**The Dissertation Committee for Colin Henry Krainin Certifies that this is the approved version of the following dissertation:**

# Commitment and Conflict

**Committee:**

Thomas Wiseman, Supervisor

Terrence Chapman

Patrick McDonald

Maxwell Stinchcombe

Dale Stahl

# Commitment and Conflict

by

**Colin Henry Krainin, B.A.; M.S.Eco.; M.S.Eco.**

**DISSERTATION**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**DOCTOR OF PHILOSOPHY**

**The University of Texas at Austin**

**December 2013**

# Dedication

This dissertation is dedicated to

John G. Krainin

Joan P. Krainin

Joseph G. Krainin

Gillen H. Krainin

Joseph M. Krainin

Loretta E. Krainin

Henry H. Platek

Vivian L. Platek

# Acknowledgements

I am deeply indebted to Tom Wiseman, Terry Chapman, and Pat McDonald for their wisdom, advice, and encouragement. This dissertation would not have been possible without them. I would also like to thank Bei Cao, Bob Powell, Teresa Schuessler, John Slinkman, Dale Stahl, Max Stinchcombe, Harrison Wagner, and Rachel Wellhausen for their very helpful comments and suggestions.

In addition to those who contributed directly to this dissertation, I would like to thank the following teachers that deeply inspired, encouraged, and challenged me at various points throughout my academic career: Charles Davis, Mark Scandling, Tao Luo, James Sandefur, Ian Gale, Tod Linafelt, Ronny Razin, and Marcin Peski.

# Commitment and Conflict

Colin Henry Krainin, Ph.D.
The University of Texas at Austin, 2013

Supervisor: Thomas Wiseman

War is an inefficient outcome and therefore states ought to prefer to bargain over areas of conflict instead of fighting. However, in the anarchy of international relations there is no actor with a monopoly of power to enforce contracts between states. States then face a commitment problem when bargaining to prevent war. This dissertation explores three models where this commitment problem can lead to war.

The first chapter presents a model that allows for shifts in the distribution of power which play out over an arbitrary number of time periods. This leads to a sufficient condition that implies war under a broader set of conditions than previously shown in the literature. This condition implies that preventive war may be caused by relatively slow, but persistent shifts in the distribution of power. As theorized in power transition theory, differential rates of economic growth can potentially cause war under this mechanism.

Relaxing the unitary actor assumption of the first chapter, the second chapter analyzes how the domestic institutional structure of countries affects the likelihood of

war.[1] We model institutional divergence by comparing an infinitely lived dictatorship to a democracy with a replaceable leader and allow a range of leader incentives within these institutional frameworks. We show that dictators, even welfare maximizing ones, may lead to war if the initial distribution of resources is highly imbalanced whereas a democracy with a forward looking electorate is always peaceful. Yet when a democratic electorate is myopic, preventive war may result. Political parties act as a mechanism to prevent this outcome.

In the third chapter, I investigate adding a third actor to the bargaining model of war. In a static setting, the model uses a notion of cooperative stability to predict balancing and bandwagoning behavior in alliance formation. When extended to a dynamic setting, changes to the system that result in alliance shifting may cause war. Additionally, alliance formation need not correspond to the static solutions, suggesting that the dynamics of power are as important as the distribution of power in alliance formation.

[1] Based on a joint project with John Slinkman.

# Table of Contents

# 1 The Persistent Shock Mechanism for Preventive War

## 1.1 Introduction

Two lines of inquiry within the international relations literature contend that differing conceptions of shifts in the distribution of power cause war. The older power transition theory argues that shifts in power due to long term changes such as industrialization lead to hegemonic challenges of the currently dominant state.[1] More recently, Fearon (1995) puts forward the "commitment problem" as a potential cause of war. Commitment problems arise when anticipated shifts in the distribution of power cause adversely impacted states to go to war in order to prevent future declines in their bargaining power. The rising state may prefer to avoid war by committing to not exercise its future bargaining power with a contract. However, since there does not exist a world government to enforce contracts between states, there is no means by which a rising state can commit to honor this contract in the future.

Powell has subsequently shown that a common mechanism underlies the various types of commitment problems outlined by Fearon. Namely, large and rapid shifts in the distribution of power cause war.[2] This result stands in direct contrast to power transition theory since as Powell points out, "differential rates in economic growth are empirically too small to account for war through this mechanism."[3] I argue that even when shifts in the distribution of power are not necessarily large and rapid, war results when they are sufficiently persistent. In order to distinguish this from Powell's large and rapid mechanism (LRM), I will refer to this mechanism throughout this chapter as the persistent shock[4] mechanism

[1] See Organski (1968); Organski and Kugler (1980)

[2] See Powell (1999); Powell (2004); Powell (2006)

[3] Powell (2004), 238

[4] The terms "shock" refers to the terminology used in economics when referring to a technological shock to economic production. This is meant to be suggestive since this mechanism is particularly helpful in thinking about the effects of expected relative change in economic growth.

(PSM). The PSM generalizes Powell's analysis with the interesting implication that power transition theory's emphasis on long term economic changes now work as a potential cause of war in a complete information bargaining model of war.

In order to focus the attention of this chapter, I concentrate on extending Powell's model of preventive war[5] to include persistent shocks to the distribution of power.[6] In this model, the PSM leads to a broader sufficient condition for war than the LRM. This is theoretically important for two reasons. First, as suggested above, two previously oppositional strands of the literature can be brought into alignment. Second, it allows for a natural extension to the many theoretical models that utilize Powell's sufficient condition for war.[7] However, the larger contribution is empirical in that the PSM allows for a more natural reading of historical cases than the LRM allows. In order to make this point, one must be clear as to under which conditions the LRM predicts preventive war.

First, it should be pointed out that "rapid" changes in expectations only satisfy the LRM if the actual expected change will, in fact, occur rapidly. In an instant, one might change one's expectations of the size of state A's GDP in 2035. However, this would still only justify war under the LRM if the rate of change in any one bargaining period was enough to satisfy the size conditions of the LRM. Only the size of the expected transition between any one period matters. Under the PSM, using the same length of bargaining period, one can predict war today based off a change in expectations of future events such as GDP in 2035. There are still requirements as to how large the current period's shift needs to be so that war is not put off, but they are not nearly as strong as the LRM implies.

This then suggests the question, what constitutes a "rapid" change in a war model? In terms of theory, it means one model period. Yet a model period is as long as the modeler

[5] See Powell (2006), 181-184

[6] Appendix B demonstrates that the results in this chapter generalize significantly. The Powell (2006) model is chosen for its simplicity in exposition.

[7] For example: Leventoglu and Slantchev (2007); Schwarz and Sonin (2008); Wolford, Reiter, and Carrubba (2011); Chapman, McDonald, and Moser (2012); Powell (2012)

dictates. However, the logic of bargaining models of war become increasingly awkward as one increases the length of a model period much beyond the length of time it takes for states to exchange bargaining offers.[8] Any extension of the bargaining period beyond this must be justified by a "time flow indivisibility" of a bargain. By this I mean that the ability of two states to divide a bargain flow is limited by the length of time necessary for the flow to arise.[9] In applying a model empirically, one is limited to the frequency of data observations. Given this, and to fix ideas, I will use a year long period of time to represent a model period. Shorter periods of time only make the PSM a more compelling substitute for the LRM since shorter periods make "large" shifts harder to justify. Since long model periods only arise from time flow indivisibilities or the inability for states to make bargaining offers or war decisions at a given point of time, when such frictions are absent or small, neither short or long periods "fix" the LRM.

In fact, the "rapid" part of the LRM is the core problem in the formulation. It is difficult to justify preventive war when bargain periods need only be long enough to accommodate an exchange of offers or a flow of benefits, but, in the same time frame, allow for large changes in the underlying state of the world. Discrete changes, such as a state acquiring nuclear weapons, are needed to make the mechanism natural in an empirical setting. Since the scale of per period shifts that are sufficient to cause war under the PSM are so much smaller than under the LRM and need not fit an arbitrary degree of "rapidness," one can much more easily and naturally justify a rational preventive war motivation in historical wars. Furthermore, the PSM focuses one's thinking not on per period shifts in the distribution of power, but on the expectation of where the distribution of power will be in the

[8] I do not mean that negotiation cannot take a long time to complete, rather that one must argue that such negotiations are delayed beyond a model period in equilibrium.

[9] That the length of time is best represented by a continuous time model is not clear. It does take a discrete amount of time to exchange offers. Despite this, perhaps a continuous time model is a good approximation of modern times. The model in this chapter is a discrete time model. A similar analysis for a continuous time model would be an interesting direction for future research.

relevant future. In the language of economics, business cycle shocks are not as important as shocks to the trend.

Empirically, the LRM seems to be most applicable when looking at the causes of war which are the results of actions taken by policy makers. Large and rapid shifts in the distribution of power on the scale implied by Powell's model are, in fact, quite dramatic changes. Examples of such a shift might be the acquisition of nuclear weapons or a sudden change in the alliance structure of the international system. Both are highly relevant events in the study of war. However, saying that such actions are the root cause of an inefficient preventive war begs the question, why did the state choose to acquire the nuclear weapons or change the alliance structure in the first place? Debs and Monteiro (*n.d.*) argues that incomplete information is necessary to generate such actions when using the LRM. Under the PSM logic, exogenous shocks may induce future equilibrium actions along these lines, but it is the initial structural shock that is the root cause of war. Furthermore, while Debs and Monteiro (*n.d.*) emphasizes incomplete information as the root cause of war in order to get around the problems of justifying large and rapid shifts, the PSM shows that incomplete information is not necessary to cause war using the logic of commitment problems.

Another case where the LRM applies is the case of first strike advantages. As highlighted in several papers,[10] first strike advantages lead to rapid and potentially large shifts in the distribution of power since as one player gives up the advantage of striking first, it immediately swings to the other player. While it is true that many wars have indeed begun with one state taking advantage of a first strike, it is difficult to imagine that such advantages are the fundamental cause of a war unless they had just come into existence, again, through some exogenous technological shock or change to the international structure. Otherwise, one must ask what has changed that made these first strike advantages critical at the instance war began? As Fearon (1995) argued, the presence of first strike advantages

[10] Examples include Powell (2006); Leventoglu and Slantchev (2007)

mostly serves to increase the likelihood of war when a more fundamental cause is present.

A case where the LRM seems not to provide a compelling theoretical justification for war is when the shift in question is of an economic nature. Yet, over the last two and half centuries, the onset of industrialization led to dramatic shifts in the distribution of power and several concurrent wars. This is the LRM's core antagonism with power transition theory, which has long argued that these economic shifts were causal for wars in this time period. The problem with economic shifts, when confronted with Powell's mechanism, is that they occur over long periods of time so that they do not fulfill the rapid part of the formulation. However, the PSM largely reconciles these two theories. What still lies in contradiction with power transition theory, is that power transition theory places an emphasis on war caused by rising challenger disrupting the place of the currently dominant state, but the PSM makes no such claims. What matters is still the logic of Fearon – war today prevents a worse peace in the future.

In order to highlight the greater explanatory power of the PSM, in a later section I will provide an illustrative example of Germany's preventive motives against Russia at the beginning of World War I. The PSM puts a greater emphasis on the long term combination of economic, financial, and military changes brewing within Russia, while looking for a shift that is large and rapid enough to satisfy the LRM is difficult. If anything, the emphasis of the LRM would be on military trends alone. In general, the LRM encourages the empiricist to focus very narrowly on short term trends and shifts, while the PSM allows for both short term and long term motivations.

For instance, consider then Secretary of Defense Richard Cheney's comments on the 1990 Iraqi invasion of Kuwait:

> [Saddam Hussein] has clearly done what he has to do to dominate OPEC, the Gulf and the Arab World. He is 40 kilometers from Saudi Arabia, and its oil

> production is only a couple of hundred kilometers away. If he doesn't take it physically, with his new wealth he will still have an impact and will be able to acquire new weapons including nuclear weapons. The problem will get worse not better.[11]

The LRM would focus one on the first part of the statement and the prevention of an anticipated shift in power if Saudi Arabia was then invaded. Yet this argument only gets part of the way there. If the second, longer term consideration were absent, US intervention does not necessarily follow. Certainly even the anticipated shock of a conquered Saudi Arabia would not increase the short term military strength of Iraq in a way that threatened the US. Even if this was the concern, it is not clear why a defensive strategy would not have been optimal. The PSM emphasizes the second concern, that with Kuwait and possibly Saudi Arabia conquered, Iraq would be able to acquire over time both economic and military strength which would adversely affect US interests. This, along with long term reputational concerns, make a far stronger case for the offensive intervention that would eventually follow.

In this way, many cases of preventive war become more natural when applying the PSM logic. Several pre-industrial wars of succession can be viewed as wars aimed to prevent the long term effects of an adverse occupant of an important throne. The wars of the later industrial revolution as wars to prevent the renegotiation of the international settlement implied by the rise of Germany, Russia, and the US. This logic is distinct from the LRM in that it need not rely on short term shifts as explanations, but instead focuses on long term trends. It is distinct from power transition theory in that there is little need to justify war in terms of a challenge to a currently dominant state. Instead, the empirical prediction is that the force emphasized in power transition theory, long term shifts in the distribution of power, are combined with the causal logic of the commitment problem of Fearon and

[11] NSC August 3, 1990, 3-4

Powell.

To fix ideas, I have framed persistent shifts in terms of economic growth. However, the way the military balance of power is altered by shifts in economic power is not well understood.[12] If it takes a long time to convert economic power into military power, wars are expected to be quick, and military spending cannot be carried out in secret, then states may be able to credibly commit to a weak military position in the future. This is because any action taken toward actualizing their economic power into military power could lead to an attack. If any of these conditions break down, then it is unclear that a state can commit to not exercising future economic power in international bargaining. The way military technology has changed, these dynamics are a particularly interesting area of investigation. A state today, in interacting with a nuclear armed state, may have a powerful commitment device in transparently not pursuing nuclear weapons. Conversely, a state on the precipice of World War I would have a difficult time convincing other states that it would not be able to exercise its future economic power in an existential conflict. Despite the simplifications of this chapter, it is clear that economic strength, political institutions, and military investment are deeply entangled.

Chadefaux (2011) has recently supplied a critique of assuming that current and future resource values are not subject to negotiation in his formulation of the commitment problem. This chapter's model continues to assume that power and power shifts are entirely exogenous and bargains have no endogenous effect on power.[13] This is a stark assumption. However, so is the assumption that power, especially shifting power, is freely transferable. The reality is somewhere in between, so that while power transfers may ameliorate the PSM, it does not eliminate it. I make the assumption that I do here in order to focus on the PSM abstracted from other considerations.

[12] Jackson and Morelli (2009) provides an analysis when the economic capacity of states are fixed.

[13] Appendix B generalizes this assumption.

## 1.2 The Model

Consider a model very similar to the complete information preventive war model in Powell (2006). Let there be two states, $S$ and $R$. These states bargain to split a pie of size one in every period $t \in \{0,1,2,...\}$. If the states agree on a bargain $x \in [0,1]$, then state $S$ receives a payoff of $x$ for that period and $R$ receives a payoff of $1-x$. Both states discount the future at the same constant rate $\delta \in (0,1)$. The total present value of the current and future flows of this pie is then

$$B = \sum_{t=0}^{\infty} \delta^t = \frac{1}{1-\delta}$$

In general bargaining problems the players often have an outside option that they can revert to if they do not like the bargaining outcome. Since the players always have this option, it can be thought of as a lower bound on their minmax value since there is no way the other player can force them to take less than this value. In war models, one thinks of this as being the war value for a state. A simple way of modeling war is as a game-ending costly lottery.[14] This value in a given period will then serve as a state's minmax value for that period, $M_{it}$. The lottery is costly since the essential motivating principle behind bargaining models of war is that war is an inefficient way of dividing a pie. This is the assumption that the cost of war is in excess to the cost of bargaining. More subtly, it is also the assumption that the decision maker is maximizing the flow of benefits received with no other relevant considerations such as future electoral success. This is a strong assumption, but I make it here in order to focus solely on the role of persistent shifts in the distribution of power.

Let $p_t \in [0,1]$ be the probability that $R$ wins a war at time $t$ while $1-p_t$ is $S$'s probability of victory. If a state wins the war, it receives the entire value of the pie from that period

[14] In the tradition of Wagner (2000), many recent war models look at war as a process. The conclusions of this chapter extend to this setting.

forward and nothing if it loses. However, when states resort to war, some of the pie is lost due to the inefficiency of war represented by the value $d \in (0,1)$, where, for reference, a value of $d = 0$ means no inefficiency to war and a value of $d = 1$ implies that the entire pie is lost in a war. This gives minmax values at $t$ of

$$M_{Rt} = \frac{p_t(1-d)}{1-\delta}$$
$$M_{St} = \frac{(1-p_t)(1-d)}{1-\delta}$$

Powell's formulation of the bargaining problem has three distinctive features. First, and most importantly, Powell does not allow states to be able to commit to a contract to divide up the entire value $B$. Instead, states can only commit to the division of the pie for a single period at a time. This is critical since if states could commit to a contract, the inefficiency of war would imply that war could not occur in equilibrium. This first assumption, when combined with the second assumption that the probability of victory in war is possibly shifting over time, is what leads to war in this model. Since states cannot commit to not exploiting future changes in their bargaining power that results from shifts in their minmax values, a state may choose to go to war in the current period in order to lock in its current minmax value. This is optimal when this value exceeds the valuation of future bargains after an adverse change in minmax or war values. The third distinctive feature is that Powell does not specify a specific bargaining protocol with which to solve his model. Instead, he finds a condition that implies war for any possible bargaining protocol.

This chapter will largely stick to these conventions, but with the following two changes. First, I allow for shifts in the distribution of power that take place over multiple periods. Second, I show that in the setup of this model, a take-it-or-leave-it offer bargaining protocol can match the relevant features of allowing for any bargaining protocol.[15] Namely, if $R$ is

[15]This is not the case in a more general setup. However, the major conclusion about persistent shifts still

the rising power and $S$ has the power to make take-it-or-leave-it offers in every period, then if war occurs under this bargaining protocol it will occur for all other bargaining protocols. I refer to this bargaining protocol, with $S$ making the take-it-or-leave-it offer, as $\mathbb{S}$.[16]

The model begins at period $t = 0$ when both states learn that $R$ will rise relatively to $S$ in terms of its war value so that $p_t$ is increasing. I assume that $p_t$ increases at a constant rate $\theta > 1$ for $T \geq 1$ periods starting in period.[17] At $T + 1$, the shift in power stops for the rest of the game. In symbols, this means that $p_{t+1} = \theta p_t$ for $t \in \{0, ..., T\}$ where $\theta > 1$ and $p_{t+1} = p_t$ for all $t \geq T + 1$. Alternatively, we have for a given $p_0 \in [0,1]$, $p_{t+1} = \theta^t p_0$ for $t \leq T$ and $\overline{p} \equiv p_t = \theta^T p_0$ for $t > T$. In order for this process to give a valid probability, $\overline{p} \leq 1$, $T$ must be restricted so that $T \leq \frac{-\log(p_0)}{\log(\theta)}$. Note that this type of shift can replicate any one period shift by letting $\theta = \frac{\overline{p}}{p_0}$ and defining $\overline{p}$ accordingly. Therefore, this method of shifting $p_t$ is a generalization of the shifts analyzed in Powell's preventive war model.[18] Despite the specificity of this shift, it should be clear that the type of shift considered here is not meant to be a realistic quantification of power shifts over time, only a simple and illustrative one. A more accurate quantification would utilize economic growth theory and need to address the conversion of economic strength into military power.[19]

I refer to the game that this model describes as $\Gamma$. $\Gamma$ is a stochastic game with two state variables. The first state variable is $p_t$ which defines the outside option values (or war values or minmax values) in each period. The second state variable is whether war has

go through in a more general setting. See Appendix B.

[16]Appendix A discusses how alternative bargaining protocols affect the analysis.

[17]Both Powell (1999) and Powell (2012) contain models with shifts in power occurring over several periods. Moreover, Powell (2012) can generate several empirically observed features of war such as persistent fighting, negotiated settlements, and recurrent fighting. However, neither demonstrates how long term shifts can imply war when shifts are not large and rapid. This is the focus of this chapter.

[18]A more direct generalization would be to let $p_t = p_{t-1} + \triangle$. In fact, an analogous analysis holds for this type of shift. I do not use this sort of shift here for two reasons. First, the exposition turns out to be more difficult since the size of shifts damp down over time in present value terms, making the sufficiency condition more complicated. Second, a multiplicative shift is more natural than an additive shift when comparing theoretical results with empirical realities since percentage change makes a clearer comparison than absolute change.

[19]Appendix B provides sufficient conditions applicable to more general types of power shifts.

occurred previously or not. In the case where war has occurred, states can take no further actions, so in the following analysis I assume that war has not occurred. If war has not occurred yet, then each model period $t$ begins with both state observing the state $p_t$, then bargaining occurs according to the bargaining protocol $\mathbb{S}$. Under bargaining protocol $\mathbb{S}$, $S$ makes a take-it-or-leave-it offer of keeping $x \in [0,1]$ such that if neither state declares war, then $S$ receives $x$ and $R$ receives $x$. After $1-x$ is determined, states and $R$ play actions $A_i \in \{a, w\}$ in a random order.[20] If $a$ is chosen, then that state accepts the bargain. If both states play $a$, then the bargain is accepted and no war occurs this period. If either state plays $w$, then the states do not receive the bargain value and a game-ending war with values as described previously occurs.[21] The next section gives results as to when war occurs in this game in a Subgame Perfect Equilibrium (SPE).

I also give an extension to all SPE of a more general set of bargaining protocols. In order to do this, I must define an infinite vector $P \in \mathbb{P}$ which assigns all periods $t$ an $x \in [0,1]$. Let $x_t(P)$ be the bargain value assigned at $t$ under $P$. $\mathbb{P}$ is the set of all such assignments.[22] A vector $P$ is called *minmax compatible at* $\tau$ if and only if $\sum_{\tau}^{\infty} x_t(P) \geq M_{S\tau}$ and $\sum_{\tau}^{\infty} (1 - x_t(P)) \geq M_{R\tau}$. $P$ is called *minmax compatible* if it is minmax compatible at $\tau$ for all $t \geq \tau$. Intuitively, if contracts are allowed and a bargain $P^{'}$ is minmax compatible at $\tau$, then a game beginning at $\tau$ is peaceful under a contract that specifies bargain $P^{'}$. Notice that since war is inefficient, there always exists a bargain $P^{'}$ that is minmax compatible at a given $\tau$. Without contracts, war is avoided only if $P^{'}$ is minmax compatible for all $t \geq \tau$ if a game begins at $\tau$. Otherwise, at least one state $S$ or $R$ optimally deviates at time $t$.

More technical readers will note that the model presented here is a particularly simple

[20] The random ordering excludes uninteresting equilibria from having to be refined away.

[21] I assume that $i$ chooses $a$ when indifferent between $a$ and $w$.

[22] Note that any SPE of an arbitrary bargaining game that satisfies our assumption an $x$ will be proposed within the model period before the war decision is made will have the property that it assigns $x$ according to some $P \in \mathbb{P}$. The SPE that corresponds to $P$ may be complex in the manner in which it assigns $x$, but it can always be summarized directly by $P$.

model of preventive war. However, there is a broader logic that motivates the characterization section of this chapter, which is derived by combining ideas from the literature on *bargaining with payoffs as you go*[23] with that of the strategic real options literature.[24] The essential idea throughout is that one can project out what the value of actions in the future will be using the logic of strategic real options. However, since bargaining is possible, the value in different periods will actually be within a range as opposed to a single point. If there are no possible values in the bargaining space preferred to the outside option of war for a state, that state will choose the action of initiating war. Given optimal actions and war decisions at future points, one can backward induct to the current period and determine whether a state optimizes by engaging in preventive war or not. This logic generalizes to many war models and not just the illustrative one presented here. Hence, the logic of the sufficient condition in this model applies to any model that might include persistent shocks.

## 1.3 The Persistent Shock Mechanism

Powell (2006) gives a sufficient condition for war at time $t = 0$ in his model which corresponds to a shock of length $T = 1$ in this model. I restate it here, adjusted for this model.

**Proposition 1.1.** *If* $\delta M_{R1} - M_{R0} > B - [M_{S0} + M_{R0}]$*, then war occurs in any SPE of the game,* $\Gamma$*, under bargaining protocol* $\mathbb{S}$ *at* $t = 0$*.*

*Proof.* Subcase of the proof in Powell (2004), pg. 238. Also follows directly from the proof of Proposition 1.2 in Appendix C. □

Powell states this as an inefficiency condition. Since bargaining is always efficient in this model, it must be that it is a sufficient condition for war. Rearranging terms, the

[23] See Haller and Holden (1990); Fernendez and Glazer (1991); Muthoo (1995); Busch and Wen (1995); Lee and Sabourian (2007); Abreu and Pearce (2007).

[24] A large literature including Fudenberg and Tirole (1985); Dutta and Rustichini (1993); Pawlina and Kort (2006).

inequality in Proposition 1.1 can be restated as

$$M_{S0} > B - \delta M_{R1} \tag{1.1}$$

That war is inefficient amounts to assuming that $M_{R0} + M_{S0} < B$. However, if $R$'s war value is rising then it is possible that (1.1) is satisfied. Using our model of war values one can quantify how large this shift is in terms of our parameters. Directly plugging in, one gets the following equation:

$$\frac{(1-p_0)(1-d)}{1-\delta} > \frac{1}{1-\delta} - \frac{\delta(\theta p_0)(1-d)}{1-\delta}$$

After some algebraic manipulation, this inequality gives us a condition on the size of $\theta$ that is necessary to cause war:

$$\theta > \frac{1}{\delta} + \frac{d}{\delta p_0(1-d)} \tag{1.2}$$

There are a few points to note from this equation. First, equation (1.2) defines Powell's LRM for this model. This is the size of a shift in the distribution of power that is necessary to cause war under this mechanism. The second is that at $\delta \to 1$, the RHS of (1.2) goes to $1 + \frac{d}{p_0(1-d)}$ which is analogous to the result in Powell (2006) with adjustments made for $\theta$ being a percentage change. The third point to notice is that the size of a $\theta$ that would satisfy (1.2) is quite large. To be concrete, consider a $t$ that represents one year. The economics literature suggests that a reasonable $\delta$ is then 0.96. Let $d = 0.2$ or a 20% inefficiency of war. Finally, let $p_0 = \frac{1}{3}$ so that $S$ starts of twice as likely to win a war as as $R$. This means that $\theta \gtrsim 1.82$, or an 82% increase in $R$'s probability of winning. As I argued in the introduction, letting $t$ represent a year is reasonable. Any shorter time period would make the size of the shift implied by the LRM even more unreasonable. Such a large shift in the relative distribution of power is highly unlikely to happen because of economic change.

It is useful to think about how economic shifts might translate into a shift in the distribution of power in this model. As a rough approximation, assume that $p$ take the functional form

$$p = \frac{GDPr}{GDPr + GDPs}$$

where $GDPr$ is the GDP of state $R$ and $GDPs$ is the GDP of state $S$. If the GDP of $S$ is twice the size of the GDP of $R$ at time 0, then this corresponds to $p_0 = \frac{1}{3}$. If we continue to let $d = 0.2$, then we know that the one period shift in $\theta$ necessary for war is 1.82. Imagine that the GDP of $S$ does not grow at all, but that $R$'s GDP grows at a spectacularly large rate of 15% a year. Then the model period would have to be longer than 8 years in length in order to justify war through the LRM. Such a long model period is extremely difficult to justify.

In order to address this issue, I now prove a condition that implies war for a broader range of $\theta$ in game $\Gamma$. I write the inequality analogously to (1.1), which is more transparent in this case.

**Lemma 1.1.** *War occurs in any SPE of the game, $\Gamma$, under bargaining protocol $\mathbb{S}$ at some $0 \leq \tau < T$ if*

$$M_{S0} > B - \delta^T M_{RT} \tag{1.3}$$

Lemma 1 follows from $S$'s outside option of war at time 0 dominating any peaceful equilibrium. $S$ receives a payoff of at least $M_{S0}$ by initiating war at time 0. The best $S$ can do in a peaceful equilibrium is to receive the entire pie in all periods until $T-1$ and then all the flow of benefits from time $T$ on less $M_{RT}$, which $R$ can secure for itself by fighting. This is $(1-\delta)^T B + \delta^T (B - M_{RT}) = B - \delta^T M_{RT}$. If $M_{S0} > B - \delta^T M_{RT}$, $S$ prefers to initiate war and no peaceful equilibrium exists.

There are two major discrepancies between this conclusion and that which Powell reaches. First, the sufficiency condition in Powell's work is more general than Lemma 1.1 since it neither relies on SPE or a specific bargaining protocol $\mathbb{S}$. Second, Lemma 1.1 does not pin down the timing of war to $t = 0$. Proposition 1.2 addresses the first point while Proposition 1.4 pins down the timing of war.

Lemma 1.2 shows that when looking for a sufficient condition for war in the game $\Gamma$ for all protocols $P \in \mathbb{P}$, it is sufficient to use bargaining protocol $\mathbb{S}$ and Subgame Perfect Equilibrium.

**Lemma 1.2.** *Let $\tau$ be defined at in Lemma 1.1. If $w$ is played at $\tau$ in any SPE of $\Gamma$ for the bargaining protocol $\mathbb{S}$, then $w$ is played at some $t$ such that $0 \leq t \leq \tau$ in $\Gamma$ for all protocols $P \in \mathbb{P}$. The converse also holds.*

Lemmas 1.1 and 1.2 directly imply that the inequality in Lemma 1.1 is sufficient for war under any bargaining protocol $P \in \mathbb{P}$.

**Proposition 1.2.** *If (1.3), then war occurs in $\Gamma$ under any bargaining protocol $P \in \mathbb{P}$ at some $\tau$ such that $0 \leq \tau < T$.*

Proposition 1.2 shows that (1.3) is sufficient for war at $\tau < T$. (1.1) is also sufficient at $\tau$. Corollary 1.1 shows that (1.3) implies (1.1) at $\tau$, but that (1.3) can hold while (1.1) fails. This means that there is a set of conditions under which (1.3) predicts war, but (1.1) fails to predict war. It is in this set of scenarios that the PSM brings new insight.

**Corollary 1.1.** *If inequality (1.1) is satisfied $\tau$, then inequality (1.3) is satisfied at $\tau$ for some $T$. The converse does not hold.*

Consider $\tau$ as defined in Proposition 1.2. At first glance it may seem strange that war occurs at $\tau$ for a shift longer than one period. One might imagine that $R$ can successfully bribe $S$ into waiting in periods $\tau < t \leq T$ such that $M_{Rt} \leq B - \delta M_{St+1}$ which may occur while $M_{S\tau} > B - \delta^T M_{RT}$ is still satisfied. This turns out to not be true.

First, consider the case where $\theta \leq \frac{1}{\delta}$. In this case, $\theta$ is so small that it is always the case that $M_{Rt} \geq \delta M_{Rt+1}$. This implies that $\overline{M}_R = \max\{M_{R0}, \delta M_{R1}, \delta^2 M_{R2}, ...\delta^T M_{RT}, ...\} = M_{R0}$. By assumption of the inefficiency of war, $M_{R0} + M_{S0} < B$. Since $M_{R0} \geq \delta^T M_{RT}$, then $M_{S0} + \delta^T M_{RT} < B$. This along with some further details outlined in the proof of Proposition 1.3, implies that $\theta > \frac{1}{\delta}$ is a necessary condition for war in $\Gamma$ when $S$ always offers a minmax compatible bargain.

**Proposition 1.3.** *If war occurs in $\Gamma$ for a bargaining protocol $P \in \mathbb{P}$ such that $P$ is minmax compatible for $R$ and is minmax compatible at $t = 0$ for $S$,*[25] *then $\theta > \frac{1}{\delta}$.*

Now consider the case where $\theta > \frac{1}{\delta}$. $S$'s value of war is decreasing relatively faster than its patience level. If $M_{S\tau} > B - \delta^{T-\tau} M_{RT}$ is also satisfied, then $R$ must capture, in present value terms, at least $\delta^{T-\tau} M_{RT}$ in any peaceful bargain starting at $\tau$. Consider a bargain $x$ at $t = \tau$, such that both $S$ and $R$ prefer not to go to war in period $t = \tau$. The most $S$ can hope to capture of the present value of the pie starting in period $t+1$ is $B - \delta^{T-(\tau+1)} M_{RT}$. Therefore, it must be that $S$ gets a payoff of $1 - x + \delta\left(B - \delta^{T-\tau-1} M_{RT}\right)$ under bargain $x$. If this is peaceful, then

$$1 - x + \delta\left(B - \delta^{T-\tau-1} M_{RT}\right) \geq M_{S\tau} > B - \delta^{T-\tau} M_{RT}$$

$$1 - x + \delta\left(B - \delta^{T-\tau-1} M_{RT}\right) > B - \delta^{T-\tau} M_{RT}$$

$$x < 1 - (1-\delta)B = 1 - \frac{1-\delta}{1-\delta} = 0$$

[25] Minmax compatibility for $S$ at $t = 0$ is necessary to make a substantive statement here. For any $d < 1$ there always exists a player with a positive value for war. Let $S$ be that player at $t = 0$. Let the bargain protocol $P'$ assign values so that $x_t(P') = 0$ for all $t$. $S$ then optimally chooses war at $t = 0$.

Therefore, in order for there to be a peaceful bargain, $x < 0$, which is a contradiction since $R$ cannot receive a negative portion of the pie in any given period.

The remaining discrepancy between Powell's sufficient condition and the condition in Proposition 1.2 is that Powell's condition pins down the timing of war. This is because no delay can occur if (1.1) is satisfied, whereas delay may occur under Proposition 1.2 since might prefer a bargain in period $t = 0$, then fight a war at some later date $t \leq T$. This can happen in this setting for two reasons. First, the smallest $\theta$ that satisfies (1.3) as $T \to \infty$ can be quite small. Namely, $\theta$ need only be greater than $\frac{1}{\delta}$. Second, since $\theta$ is a percentage change so that when $p_0$ is close to zero, the absolute change of a shift can be quite small. In fact, each additional percentage increase always causes a larger absolute change.

However, the fact that delay may occur provides no substantive problem, in fact it illustrates one way information is lost by using too coarse of a model period.[26] Furthermore, it conforms to the empirical expectation of observing some power convergence before a preventive war. Proposition 1.4 pins down a condition that along with (1.3), is sufficient for war to occur at any time $t$ that is reached. Hence, Proposition 1.4 fully generalizes Powell (2006).

**Proposition 1.4.** *War occurs in $\Gamma$ under any bargaining protocol $P \in \mathbb{P}$ at $t$ if (1.3) and*

$$\theta > \frac{1}{\delta} + \frac{1}{p_t}\left[\frac{d\,(1-\delta)}{\delta\,(1-d)}\right] \tag{1.4}$$

Using (1.4), one can identify a particular period $t$ at which war must occur. (1.4) implies that $S$ is more willing to put off war if $S$ is less patient, war is costlier, and the absolute increase in $R$'s war making ability is smaller because it begins from a lower base. Initially, if (1.3) is satisfied, but $p_0$ is too low for $\theta$ to satisfy (1.4), $S$ is willing to delay war since the current value of the bargain is greater than the increased cost of eventually fighting. As

[26] Too coarse of a a model period can also cause the opposite problem. A model may fail to predict war by underestimating the significance of a power shift that slows over time.

$p_t$ grows over time, the increase in the cost of war gets proportionally higher until (1.4) is satisfied. In other words, satisfying (1.3) implies that war must occur eventually, also satisfying (1.4) implies that *R*'s increase in war making ability outweighs *R*'s ability to pay off *S* in the current period. Since, by (1.3), war occurs eventually, this increase in war making ability will indeed be used. Hence, *S* holds off war so long as war today is less valuable than the entire current period bargain plus war tomorrow. Interestingly, while the context that makes war inevitable is derived from the commitment problem over the future bargain emphasized in (1.3), in the face of inevitable war, (1.4) predicts that leaders who begin preventive wars would make their decisions in order to prevent a worse future war. As Chancellor Bethmann-Hollweg responded to the suggestion that a diplomatic accommodation may have been made with Russia in 1914, "Who can say? But if war had come about later, Russia would have been in a better position. Where would we have been then?"[27] [28]

The intuition for Proposition 1.4 can also be seen in terms of two extreme bargaining protocols. Lemma 1.1 showed that if *S* has all the bargaining power, then the changing value of *R*'s war value is the key to determining if war will occur eventually of not. This is because if a shift in power ends peacefully, *S* would be able to capture of all of the inefficiency of war. *R* will only capture its war value. On the other hand, Appendix A shows that if *R* makes the bargaining proposals instead of *S*, then the satisfaction of inequality (1.4) alone is enough to imply war.[29] This is because once a shift in power ends peacefully, *R* instead of *S* will capture all of the inefficiency of war. Hence, *S* will only receive its war value in the future. Now it is *S*'s war value that is critical in determining whether war occurs or not. In Proposition 1.4, since inequality (1.3) is satisfied, then regardless of the

[27] Quoted from Copeland (2000), 84

[28] Although outside this model, the actual reality of the timing of war is more complex. A richer model might relax the unitary actor assumption and introduce a leader dependent on public support for war initiation. This leader would have an incentive to time wars around crises if public support for war peaks after a crisis and then fades.

[29] This result is also implicit in Powell (2012).

bargaining protocol used, war will occur at some point $\tau < T$. At $\tau$, $S$ receives its war value, so it is *as if* $R$ has proposal power at $\tau$. Therefore, in periods leading up to $\tau$, the condition for war when $R$ has proposal power becomes decisive. Hence, (1.3) and (1.4) combine to determine when war will occur for any bargaining protocol.

Using Proposition 1.4, I can now define the PSM for $\Gamma$ under bargaining protocol $\mathbb{S}$. The PSM can be formulated in two ways. In the first formulation, shifts are allowed to be sufficiently persistent (large enough $T$) that (1.3) is satisfied. Defining $t$ as in Proposition 1.4, I let $t = 0$ for convenience.

**Proposition 1.5.** *(PSM 1) For a sufficiently persistent shock $\theta$, war occurs in any SPE of the game, $\Gamma$, under bargaining protocol $\mathbb{S}$ at $t = 0$ if and only if $\theta$ is large enough to satisfy (1.4) at $t = 0$.*

Depending on the initial condition for $p$ and parameter $d$, PSM 1 can imply war in a range that is significantly larger than the LRM implies. This range can be defined as the space of $\theta$s that satisfy (1.4) but not (1.2) which is exactly $[\frac{1}{\delta} + \frac{1}{p_0}\left[\frac{d(1-\delta)}{\delta(1-d)}\right], \frac{1}{\delta} + \frac{d}{\delta p_0(1-d)}]$. The size of this space is $\frac{d}{p_0(1-d)}$. In the earlier example where $p_0 = \frac{1}{3}$, $\delta = 0.96$, and $d = 0.2$, the LRM predicted war when $\theta \gtrsim 1.82$. Using PSM 1, war is predicted whenever $\theta \gtrsim 1.07$ or a 7% shift in $p$. In fact, under $\mathbb{S}$, the PSM is both necessary and sufficient. Therefore, there can be no improvement on the war predictions of the PSM for this model.

Alternatively, if (1.4) is assumed to hold, PSM II gives the condition for how large a shock $\theta$ must be for a given persistence level ($T$) in order to cause war.

**Proposition 1.6.** *(PSM 2) For $\theta$ large enough to satisfy* (1.4)*, war occurs in any SPE of the game, $\Gamma$, under bargaining protocol $\mathbb{S}$ at $t = 0$ for a given shock length $T$ if and only if*

$$\theta > \frac{1}{\delta}\left(1 + \frac{d}{p_0(1-d)}\right)^{\frac{1}{T}} \tag{1.5}$$

This formulation lets one see exactly how persistent a shift level must be to imply a particular $\theta$ threshold. The table below gives the minimum $\theta$ values that cause war under several different values of $T$.

Table 1: PSM 2 under $\mathbb{S}$: $p_0 = \frac{1}{3}$, $\delta = 0.96$, $d = 0.2$

| $T$ | 1 | 2 | 3 | 4 | 6 | 8 | 10 | 12 | 14 |
|---|---|---|---|---|---|---|---|---|---|
| Min $\theta$ | 1.82 | 1.38 | 1.26 | 1.20 | 1.14 | 1.12 | 1.10 | 1.09 | 1.08 |

For $T > 14$, the minimum $\theta$ starts to approach the level dictated by (1.4). Strikingly, if one thinks of a decision maker thinking only a limited amount of periods into the future, the size of the anticipated shift in the distribution of power needed to imply war falls dramatically with respect to only looking forward one period as in the LRM. If decision makers are modeled in this way, the theory of preventive war caused by anticipated shifts in the distribution of power becomes much more robust.

## 1.4 World War I

Many historians and political scientists have argued that Germany was at least partly induced into beginning World War I by a preventive war motivation against rising Russian power.[30] In this section, I will see if one can make sense of this argument only using the model in this chapter. A large caveat must be made that the model in this chapter dramatically oversimplifies the situation in 1914.[31] However, this exercise highlights the analytical power of the PSM in comparison with the LRM. Furthermore, it gives insight into the types of methods that might be used in the structural quantification of formal models of war in order to understand the causes of historical wars.

[30] This was also the opinion of many of the German decision makers. For instance, the German Chancellor Bethmann stated in 1917, “Yes, by God, in a way it was a preventive war” (quoted from Mombauer (2001), 189). Also see Copeland (2000) for a detailed argument that German decision makers wanted and instigated war with Russia despite the risks of world war.

[31] The fact that I am ignoring multilateral interactions with Britain, France, and Austria-Hungary is unrealistic. Chapter 3 addresses embedding the bargaining model of war in a multi-state setting.

In order to fit $\Gamma$ to the pre-World War I situation, values must be estimated for five parameters: $\delta$, $d$, $p_0$, $\theta$, and $T$. Additionally, one must decide on a bargaining protocol. By the results of Proposition 1.2, using $\mathbb{S}$ makes the most compelling argument since if war occurs under $\mathbb{S}$, it would occur for any possible bargaining protocol. For $\delta$, I will continue to assume bargaining period of one year. Therefore I will set $\delta = 0.96$. I appeal to my arguments in the introduction that it is difficult to justify a longer period length. Shorter periods only make the following argument stronger.

One argument for why Germany conducted a preventive war might be that it was attempting to prevent the completion of the Russian "Great Program" that would have increased Russia's standing peacetime army by 40% in 1917.[32] If one agrees that periods of length greater than a year are not compatible with the LRM, then this argument is already inherently using the PSM logic since it implies a shock of at least three years was the cause of war. In order to use the LRM with this parametrization, one would need to argue that the anticipated progress made by Russia between 1914 and 1915 alone could justify a preventive war. This highlights the flexibility that the PSM grants in relaxing "rapidity" is natural in historical arguments that cite anticipated events not in the immediate future.

While this build up in military prowess was an important aspect of Germany's preventive motivations, the PSM points us more towards slower economic shifts that were occurring at the time. In order to illustrate this, I will look at the most generous parametrization of only relying on military shifts to justify Germany's preventive war. I will then compare this to including economic shifts in the argument.

A typical functional form for $p_0$, where $p_0$ is taken to be Russia's probability of victory at $t = 0$, takes an underlying state variable in each country labeled $K_{i0}$ and calculates $p_0$

[32] McDonald (2011), 1106

according to the following formulation:

$$p_0 = \frac{K_{R0}^{\lambda}}{K_{R0}^{\lambda} + K_{G0}^{\lambda}}$$

$\lambda$ is a parameter that measures the curvature of the probability function. For simplicity, I let $\lambda = 1$, but the results obtained work for other reasonable values as well. For $K_{i0}$, I would ideally use a value that well approximates military effectiveness. Some appropriate mix of military size, military expenditure, population, and economic power would most accurately represent this. A precise argument as to how to do that is beyond the scope of this chapter. Instead, I simply use iron/steel production in each country in 1913 as a proxy for $K_{i0}$.[33] This approximates the industrial might that could be brought to bear in a war at the time. It is a good rough measure for my purposes since it skews the probability of victory the most towards German victory of all the reasonable measures available. This makes for the smallest possible $p_0$, which means that it is the most difficult case to satisfy the PSM conditions on war for Germany since the RHS of (1.4) and (1.5) are decreasing in $p_0$. Using this measure I get a value of $p_0 = 0.22$.[34]

In order to estimate $d$, one must know how much inefficiency loss would occur in a war. The relevant valuation for $d$ would have been the estimation of cost by the decision makers in Germany and Russia. Most likely, decision makers would not have been able to accurately predict just how costly World War I would prove to be. Still, a war between major powers was seen as very destructive. Since it is difficult to put a precise number on $d$ without deep historical research, I will simply solve for the minimum $\theta$ that satisfies a given level of inefficiency loss for persistence level $T$. The higher the level of $d$ that results in war, the stronger the argument that a shift in the distribution of power towards Russia

[33]Data from the Correlates of War data set.

[34]A low $p_0$ (probability of victory for Russia) also seems justified by the poor ex-post Russian performance in World War I, even with a large amount of German forces devoted to fighting France and Britain.

caused a preventive war in 1914. Table 2 presents these results.

Table 2: Minimum $\theta$ for which the PSM Predicts Preventive War in 1914

| *d*/*T* | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 0 | 12 |
|---|---|---|---|---|---|---|---|---|---|
| 1% | 1.09 | 1.07 | 1.06 | 1.05 | 1.05 | 1.05 | 1.05 | 1.05 | 1.05 |
| 2% | 1.14 | 1.09 | 1.07 | 1.07 | 1.06 | 1.06 | 1.05 | 1.05 | 1.05 |
| 5% | 1.29 | 1.16 | 1.12 | 1.10 | 1.09 | 1.08 | 1.07 | 1.06 | 1.06 |
| 10% | 1.57 | 1.28 | 1.19 | 1.15 | 1.13 | 1.12 | 1.10 | 1.09 | 1.08 |
| 15% | 1.88 | 1.40 | 1.27 | 1.21 | 1.17 | 1.15 | 1.12 | 1.10 | 1.09 |

How large of a $\theta$ would the Great Program alone have predicted? Over 4 years the increase in military size was expected to be 40%. The German military was also increasing in strength, but not nearly as quickly. It is therefore hard to justify a shift in $p$ over 4 years of more than 40% using military changes alone.[35] A four year change of 40% means $\theta^4 = 1.4$ or $\theta \approx 1.09$. This means that the Great Program alone would have justified war as a persistent shock only if the inefficiency loss was seen as being below the 5% level. This result is suggestive that an argument that combined the military story with one of longer term economic shifts would better justify the preventive war motive. A shift of $\theta = 1.09$ that was persistent at the $T = 12$ level would have justified a preventive war at the 15% inefficiency level. In order to have used the LRM to predict war at the 15% level, one would have needed an anticipated one period shift in $p$ of 88%. Given that in 1913 Russia had 2.5 times the population of Germany, but only 28% of the iron/steel production indicates that there was a lot of room for convergence of economic power now that Russia had begun to reform its economic institutions.

There is still a great deal to be sorted out in the modeling of an historical case like this. In the analysis of McDonald (2011), emphasis is placed on a state's ability to translate economic resources into military resources given domestic political constraints. This is a

[35] A relatively faster increase in military quality might make the rate more rapid. However, since I am assuming no increase in German size or quality this is at least somewhat counteracted. The four year change would have had to have been on a level of a 75% increase in order to justify a $\theta$ of 1.15 and hence satisfy $d$ at the 10% inefficiency loss level. Therefore, this analysis is robust to small changes in these estimates.

critical area to understand in the translation of economic shifts into shifts in military power. However, by building quantifiable models such as the one used here, one can begin to put a larger amount of rigor in the size of effects resulting from possible reasons for historical war, as opposed to directional arguments or reduced form correlations.

## 1.5 Conclusion

From the sufficient conditions proved in this chapter comes the insight that preventive war may occur under a broader set of circumstances than just large and rapid shifts in the distribution of power. Namely, small, slow, and persistent shifts in the distribution of power can cause preventive war. This means that war may often be caused not by the anticipation of an adverse shift in the distribution of power in the immediate future, but by changes in expectations today about future events, even if these events are quite distant.

Predicting preventive war under a broader set of conditions allows for a greater amount of flexibility in applying formal models to historical cases. In some sense, the theory is catching up to what has already been done for a long time in historical analysis. Now, though, historical analyses that make an argument that a particular war was primarily preventive in nature can set their arguments within the PSM framework. This allows for quantification and a much more compelling case that the size of a given preventive motive was sufficient to cause war.

There are several possible theoretical extensions to the model in this chapter. In many ways, the model presented here is the simplest possible that utilizes the logic of the PSM. Future models might interact the PSM with domestic considerations, multilateral effects, incomplete information, and especially more general processes for shifts in the distribution of power. Of particular interest is how economic shifts, political constraints, and strategic choices of military spending interact. A more detailed analysis of these factors could potentially explain how long term economic shifts combine with faster shifts in the political and military realms to cause preventive war.

# 2 A Theory of Democracy, Dictatorship and Political Parties in International Bargaining[36]

## 2.1 Introduction

Bargaining models of war typically assume that countries[37] themselves have well-defined preferences and streamlined decision making processes. This unitary actor assumption is a useful fiction, one well-founded within the international relations literature and providing deep insight into how countries behave. Yet relaxing the assumption reveals a complex system of strategic interactions underlying a country's choice of representative on the international stage. This chapter looks at how a leader's characteristics and the nature of the domestic institutions involved in choosing a leader are critical to the bargaining process between countries and, ultimately, to whether wars occur. We drop the unitary actor assumption, making the nature of a leader a strategic choice for one country's populace. The key notion is that a leader's incentives may not align with the incentives of that populace and we vary the leader's incentives using a notion of political bias first defined in Jackson and Morelli (2007). This bias, which causes the leader to evaluate the relative costs and benefits of war differently than the population he represents, can be thought of as either modeling the leader's own preferences or as a reduced form model of the political institutions that shape the leader's incentives. We also vary the strategic setting by varying the institutions under which the choice of leader occurs. We find that differences in these institutions governing choice of a leader result in vastly disparate bargaining and war outcomes.

We focus on two institutional structures: one, which we call democracy, allows an

[36]This chapter is based on a joint project with John Slinkman.

[37]Since we will be using the concept of a state variable, we refer to the political unit of a state as a country. There is, in fact, nothing in our analysis that does not generalize to any bargaining between groups which choose representatives to bargain in their stead if those two groups have a "war-like" outside option.

electorate to replace its leader in each period. The other, which we call dictatorship, allows for the leader to be selected at most once at the outset of the game. Our model implies that democracies are peaceful but exceptionally greedy in the long run,[38] capturing the entirety of the bargain in the limit. Since we vary institutional structure on only one side of the bargain, we cannot comment directly on the implications of our model for the democratic peace literature. However, the incentives that make our democracies greedy do cast doubt on the interpretations about why democracies do not go to war with each other: it may not be out of mutual respect for each others' institutions, but rather that the electorate, which is inherently peaceful in our model, has more finely grained control over the period by period bargain than under other institutional structures. This result does not hold if we consider democracies that are not perfectly forward looking, as would seem to be the norm in a world with many frictions not present in our baseline model. Myopic electorates in a democratic country can cause preventive war since they fail to account for their own commitment problem while increasing their leader's bias too quickly.

Interestingly, political parties can serve to untangle this problem. By presenting voters with nominees from a fixed menu of profiles, the electorate cannot optimize over leader qualities individually. We present a simple model of this where the electorate values a domestic issue that favors one party or another stochastically in a given election. When this is the case, the electorate chooses a leader who may not be an optimal representative on the international stage. This serves as a commitment device to not increase bargain demands too much in the future. When the probability of a non-optimal leader is high enough, peace results. The foreign policy restraint implied by counteracting political objectives is what makes a democracy peaceful in our model. In that sense, this specification of our model can be viewed as a formalization of an aspect of Republican Security Theory (see Deudney (2007)).

[38] One may even call them "peacefully imperialistic."

Unlike the case of democratic institutions, various outcomes can occur in those that are dictatorial,[39] depending on the initial distribution of the resource being bargained over. When countries are on a relatively even initial footing, peace results. However, the optimal choice of dictator for countries with a relatively small share of the resource at the outset – by which we mean the choice of leader that maximizes the expected utility of a representative consumer in that country – invariably leads to war after some delay, during which the country receives concessions from its bargaining partner. Despite the delay, war occurs in these cases when the country that was weak to begin with is still relatively weak. That is, the electorate of relatively weak countries prefer to go to war at some point in the future despite the fact that this electorate faces the full inefficiency cost of war. This is because, starting with so little, they prefer rapid short run gains at the bargaining table, despite the fact that these gains eventually increase the value of war for their leader until he makes unreasonable demands that cause war. This result is thus an alternative take on the classic commitment problem in the bargaining model of war, the crucial difference being that in our model it is not sudden rapid shifts of power that lead to immediate war, but the expectation of a persistent transfer of power if preventive war is not initiated.

A particularly striking historical example of a weak country choosing a warlike dictator – one with a quite destructive eventual outcome – is that of Germany in 1933. In the years before World War II, Germany renegotiated the international bargain dramatically in its favor. However, as power accrued to Germany, Hitler increased his demands in proportion to his new power, leading invariably to war. It is striking how closely the dynamics of our model match this historical behavior.

[39]We use the term "dictatorship" to refer to actual dictatorships as well as democracies where political incentives never vary. If a leader is so constrained by the nature of her office that incentives never change even with changes in the identity of the leader, this system effectively operates as a dictatorship in our model. In this case, the choice of institutional incentives occurs only in the beginning of a game.

### 2.1.1 Related literature

Our model incorporates three interconnected bargaining frictions. The first friction is that bargains endogenously affect a country's ability to go to war and hence its future bargaining power. Fearon (1996) provides the first formal analysis of this friction in a discrete time setting. In this model the transfer of power inherent in bargaining can lead to shifts in power and hence war through the commitment problem. Schwartz and Sonin (2008) analyzes a continuous time model with similarly endogenous bargaining power. Chadefaux (2011) also provides a model where bargains can affect power. In his model, countries are allowed to bargain over power and utility separately so that the possible transfer of power does not act to exacerbate the commitment problem in bargaining, but instead as a means to avoid it. Our model is closer to Fearon (1996) where countries must bargain over an issue that simultaneously affects both utility and power. We agree with Chadefaux (2011) that there are certain issues that have a larger impact on either a country's utility or power. For instance, the transfer of naval ships may affect power much more than utility. However, there do exist many issues in which the effect on a country's utility and power, such as territorial possessions, are inextricably entangled. Our model can be viewed as isolating only those issues.

Second, the model includes domestic political considerations that translate into an agency problem. The type of agency problem we consider follows Jackson and Morelli (2007). They consider an one shot model where bargaining countries choose a leader with political bias to represent them in negotiations. The leader's bias arises from the division of the bargain she receives versus the division of the potential spoils she would receive in war. We take a slight modification of this game as our stage game in a stochastic game context.

Finally, the ability of a country to commit to a leader or a policy for electing leaders plays a critical role in determining bargaining and war outcomes. This builds on the "commitment problem" literature that runs through conceptualization in Fearon (1995) to the inefficiency condition in Powell (2004) and Powell (2006) and the broader sufficient conditions for war in Chapter 1 and Appendix B.[40] In this chapter, perfect commitment to a dictator can in fact lead to war while a perfectly forward looking democracy can avoid war even without commitment.

By analyzing the optimal selection of leaders in context of international bargaining and war, this chapter draws insight from two series of papers: Wolford (2007, 2012a,b) and McGallivray and Smith (2000, 2004, 2005, 2006). The Wolford papers take place in a finite extended form game with an emphasis placed on the strategic behavior of the leader. Our model utilizes a stochastic game context to discover the long run implications of leader selection. The emphasis here is placed on the strategic considerations of the electorate in choosing a leader or domestic institution through which a leader is chosen. The McGillivray and Smith papers focus on achieving cooperation in a repeated or stochastic game setting when countries can utilize punishment specific to leader selection. Our model differs in that it incorporates the bargaining model of war into the leader selection setting while avoiding looking at leader specific punishments for purposes of parsimony. Our focus allows us to address the long term impact of leader selection as well as different domestic institutions on bargaining and war outcomes. Adding more complexity to the leader's strategic situation, as in Wolford or McGillivray and Smith, would be an interesting extension to the current model.

The bargaining theory foundations of this chapter derives from the *bargaining with payoffs as you go* literature as well as the bargaining with behavioral types literature in economics. The former was originally formulated in analyzing wage strikes. Prominent ex-

[40]These conditions all result from the analysis of shifting outside option values. The theory of bargaining with outside options in a context similar to bargaining models of war is developed in Muthoo (1995b). Particularly relevant for this chapter is Compte and Jehiel (2002) which looks at the interactions between outside options and the possibility of behavioral types.

amples include Haller and Holden (1990), Fernendez and Glazer (1991), Muthoo (1995a), Busch and Wen (1995), and Lee and Sabourian (2007). The latter was analyzed in a great deal of generality by Abreu and Gul (2000). A combination of the *bargaining with payoffs as you go* in a repeated game context with behavioral types is found in Abreu and Pearce (2007). In this chapter, we use a simple stochastic game formulation. Although we are not aware of a paper that generalizes the non-cooperative foundations of behavioral types à la Abreu and Gul (2000) to a stochastic game setting, we are informed by Abreu and Pearce (2011) which sets Nash (1953) style perturbations in a stochastic bargaining game.

The rest of the chapter is organized as follows: Section 2.2 presents the model. Section 2.3 characterizes the game under the institutional framework of a dictatorship. Section 2.4 characterizes the democratic outcomes. Section 2.5 concludes. Appendix C.2 contains all the proofs.

## 2.2 Model

Consider two countries, $H$ome and $F$oreign, who bargain over the division of a stream of surplus where the surplus endogenously affects the distribution of power between the two countries. Formally, we will say that $H$ and $F$ must bargain over a pie of size 1 at each period $t \in \{0, 1, \ldots\}$.[41] This could represent the year to year negotiation over foreign aid, territory, or gains from trade. Critically, $H$ and $F$ do not bargain directly with each other; instead, $H$ chooses a leader, $L$, to bargain in its stead.[42] Potential leaders differ in their *bias* towards war, defined formally below.

There are two aspects of $H$'s choice of leader that are of interest: the first is the institutional structure by which $H$ chooses its leaders; the second is $H$'s actual choice given the

[41]The normalization of the pie to size 1 is without loss of generality; the fact that the size of the pie is fixed is not.

[42]Note that we model one side's choice of leader. A possibly interesting extension is to allow $F$ flexibility in choosing its leader as well.

institutional structure. The majority of this chapter focuses on two extreme models of institutional structure. The first we refer to as a choice *with commitment* in which $H$ chooses a leader once and for all at time 0. The second we call the choice *without commitment* where $H$ chooses a new leader in each period. We identify the choice of a leader with commitment with the institutional structure of a dictatorship and the choice of a leader without commitment with the institutional structure of a democracy.

### 2.2.1 Stage game timing

The timing of the stage game in period $t$ is as follows:

1. Countries observe the status quo division $y_t$.
2. *In the game without commitment, H chooses a new leader.*
3. $F$ proposes a new division of the pie. Denote $H$'s share of this division by $x_t$.
4. $L$ accepts or rejects the offer.
   (a) If $L$ accepts, all players receive their spot payoffs and the game advances to the next period with $y_{t+1} \equiv x_t$.
   (b) If $L$ rejects, game ending war occurs, with players receiving the value of their outside option.

We formally define the action sets for the stage game in Appendix C.2.1. In particular, $H$'s choice of leader is really a choice of the bias of the leader representing it, the introduction of which we delay until Section 2.2.4.

Note that $F$ has no way to explicitly declare war. This is without loss of generality, as $F$ will always be able to make an unreasonable offer, forcing war. Note also that the choice of stage game bargaining protocol at first seems extreme: because $F$ makes a take-it-or-leave-it offer to $L$, $F$ has as all of the bargaining power. Though a strong assumption, this is justifiable in two ways.[43] First, because the source of war in this model will come from $H$'s choice of leader, by granting $F$ as much flexibility in choosing the bargain as possible, conditions for war are as strict as possible. In this sense, our results concerning the occurrence of war can be seen as providing sufficient conditions for war in models with alternative specifications of the bargaining protocol. Second, though alternative bargaining protocols alter the outcome specifics, the qualitative characteristics of the model remain unchanged for a large class of bargaining protocols.

### 2.2.2 War

We model war as a game-ending, state-dependent costly lottery. Let $p(y)$ be the probability that $H$ wins a war when the status quo is $y$. Let $1-p(y)$ denote be the probability that $F$ wins. We require that $p(y)$ be strictly increasing, weakly concave, and differentiable. The simplest such functional form is the linear probability function, $p(y) = y$. War destroys a fraction $C$ of each country's share of the status quo. The winner of the war also receives a fraction $G$ of the loser's pre-war share of the status quo. Thus if war occurs when the status quo is $y$ and $H$ wins, the final division of the surplus is $(1-C)y+G(1-y)$ for $H$ and $(1-C-G)y$ for $F$. We assume that $C+G \leq 1$ and refer to $C$ as the inefficiency of war.

### 2.2.3 Stage game payoffs for $H$ and $F$

Let $L$, $H$, and $F$ have a common discount factor $\delta \in (0,1)$. If $L$ agrees to an offer $x$, the stage game payoffs for $H$ and $F$ are their respective shares of that offer: $x$ and $1-x$. If $L$ rejects the offer, $H$ and $F$ receive the expected discounted value of the final division given

[43] See Chapter 1 for a more general discussion of these arguments.

the status quo. Call these values $M_H(y)$ and $M_F(y)$. Then

$$M_H(y) \quad = \quad \frac{((1-C)y+G(1-y))p(y)+(1-C-G)y(1-p(y))}{1-\delta}$$

and

$$M_F(y) \quad = \quad \frac{(1-C-G)(1-y)p(y)+((1-C)(1-y)+Gy)\,(1-p(y))}{1-\delta}.$$

### 2.2.4 Leader bias and stage game payoffs for $L$

We model a leader's divergent incentives for war using the notion of political bias found in Jackson and Morelli (2007). Formally, this means that the leader's share of peaceful bargains may differ from their share of the spoils of war. Let $a$ represent a leader's share of $H$'s share of a peaceful bargain and let $a'$ represent their share of the spoils of war. Then a leader's stage game payoff when agreeing to an offer $x$ is $ax$. $L$'s payoff to war is

$$\frac{(a(1-C)y+a'G(1-y))p(y)+a(1-C-G)y(1-p(y))}{1-\delta}.$$

While it is conceptually useful to think about the leader's bias in terms of differing shares of peaceful bargains and spoils of war, all that matters is the ratio of the two. Thus we scale the leader's payoffs by $1/a$ so that their payoff to agreement is $x$ and their expected value of war is

$$M_L(y,B) \quad = \quad \frac{((1-C)y+GB(1-y))p(y)+(1-C-G)y(1-p(y))}{1-\delta}$$

where $B \equiv a'/a$. We refer to $B$ as the leader's bias. We assume throughout that the leader only receives utility whilst he is in power.

There are two ways to think about what bias represents in our model. One is that it captures some individual characteristics of the leader in question. For instance, leaders with strong connections to the international sector of the economy may evaluate the relative costs and benefits of war differently than those with a stronger domestic base. In this sense, $H$'s choice of leader is actually a choice amongst leaders with differing platforms. The second way of interpreting bias is that it represents an institutional mandate for rewarding leaders in times of peace and in times of war. Countries that give awards for distinguished military service or allow their leaders the first pick of the spoils of war cause their leaders to evaluate the relative costs and benefits of war differently than those that have enforceable, high legal standards governing their leaders actions in war. In this sense, $H$'s choice of bias is an act of institutional design rather than a choice amongst individual leaders.

It is important to note that if $B = 1$, the leader's incentives are the same as $H$'s. As $B$ increases above 1, the $L$'s value to war rises relative to that of peaceful bargains and we will say that such a leader is *biased towards war*. Leaders with $B < 1$ are *biased against war*. Leaders with $B = 1$ are *unbiased*. We summarize some important characteristics of the leader's minmax value in the following lemma.

**Lemma 2.1** (Characteristics of war values)**.** *Given the assumptions on $\delta$, $C$, $G$, and $p(\cdot)$, the following are true:*

1. *$M_L(y,B)$ is concave in $y$ if $B > 1$;*

2. *For all $y$, there is $B^\dagger(y)$ such that*

$$M_L(y, B^\dagger(y)) + M_F(y) = \frac{1}{1-\delta}; \quad \textit{and}$$

3. *There is a unique $B^\dagger$ such that $B^\dagger = \arg\min_y B^\dagger(y)$. Furthermore, $B^\dagger > 1$.*

### 2.2.5 Choice of leader

We model $H$'s choice of leader as a choice of bias. For the majority of the this chapter, we do not restrict the available range of biases beyond the natural requirement that $B \geq 0$. In Section 2.4.4, we consider a restricted range of biases.

In the game with commitment, $H$ chooses a bias $B$ at time 0 and this bias is fixed for the rest of the game. In the game without commitment, $H$ chooses a new bias $B_t$ after observing the status quo and before $F$ makes its offer.

### 2.2.6 The dynamic game

Formally, this model defines a dynamic game where the underlying state variable is $y_t$, the status quo in time $t$.

Strategy profiles in the dynamic game imply sequences of offers possibly terminating in war.[44] Thus we can identify a strategy profile $\sigma$ with a sequence of offers $\{x_0, x_1, \ldots, x_n\}$ and, in the game without commitment, a sequence of biases $\{B_0, B_1, \ldots, B_n\}$. Here $n < \infty$ implies that war occurs in the $(n+1)^{\text{th}}$ period and $n = \infty$ implies that war never occurs. We will define payoffs in dynamic game both in terms of strategy profiles and the implied sequence of offers. If $\sigma$ implies that war never occurs, we have

$$V_H(\sigma) = V_H(x_0, x_1, \ldots) = \sum_{t=0}^{\infty} \delta_H^t x_t$$

$$V_F(\sigma) = V_F(x_0, x_1, \ldots) = \sum_{t=0}^{\infty} \delta_F^t (1 - x_t)$$

$$V_L(\sigma; B) = V_L(x_0, x_1, \ldots; B) = \sum_{t=0}^{\infty} \delta_L^t x_t.$$

If $\sigma$ implies that war occurs in period $n$, so that war occurs when the status quo is $x_{n-1}$, we

[44] See Appendix C.2 for a formal definition of the strategy space.

have

$$V_H(\sigma) = V_H(x_0, x_1, \ldots, x_{n-1}) = \sum_{t=0}^{n-1} \delta_H^t x_t + \delta_H^n M_H(x_{n-1})$$
$$V_F(\sigma) = V_F(x_0, x_1, \ldots, x_{n-1}) = \sum_{t=0}^{n-1} \delta_F^t (1 - x_t) + \delta_F^n M_F(x_{n-1})$$
$$V_L(\sigma; B) = V_L(x_0, x_1, \ldots, x_{n-1}; B) = \sum_{t=0}^{n-1} \delta_L^t x_t + \delta_L^n M_L(x_{n-1}, B).$$

### 2.2.7 Equilibrium concept

The value functions in the previous section can be reformulated recursively. $y$ represents the current division of the pie and $x$ is $F$'s offer to $L$ in the current period. $y$ is the only payoff relevant state other than an indicator of whether war has occurred previously or not. Since the game is trivial if war has occurred, we suppress such an indicator assuming throughout that we are in a state where war has not occurred previously. This leads to the following recursive value functions when the strategy profile is $\sigma$

$$V_H(y, \sigma) = \begin{cases} M_H(y) & \text{if } x \text{ implies war under } \sigma \\ x + \delta V_H(x, \sigma) & \text{otherwise} \end{cases}$$
$$V_F(y, \sigma) = \begin{cases} M_F(y) & \text{if } x \text{ implies war under } \sigma \\ 1 - x + \delta V_F(x, \sigma) & \text{otherwise} \end{cases}$$
$$V_L(y, \sigma; B) = \begin{cases} M_L(y, B) & \text{if } x \text{ implies war under } \sigma \\ x + \delta V_L(x, \sigma; B) & \text{otherwise} \end{cases}$$

In order to take advantage of this recursive structure, we assume that the players follow Markov strategies. Markov strategies depend only on the payoff relevant state; actions played in the past affect the strategy only through their effect on the state and player's cannot explicitly condition on past actions. We view this as a reasonable restriction for an environment in which players are not able to make commitments about future play. We thus focus throughout on Markov Perfect Equilibrium (MPE) which is a Subgame Perfect Equilibrium (SPE) such that players select Markov strategies.

## 2.3 Bias With Commitment: Dictatorship

In this section, we characterize outcomes in the game with commitment. A dictator might arise under two general scenarios: In the first case, the dictator is not chosen by the population of $H$ and thus may have any bias. In the second case, the dictator is chosen by the population of $H$ to maximize $H$'s utility. Section 2.3.1 characterizes equilibrium for a leader with a fixed but arbitrary level of bias. Section 2.3.2 discusses the equilibrium when the leader's bias is chosen by $H$.

### 2.3.1 Non-optimal Bias

The principle result here is Proposition 2.1.

**Proposition 2.1.** *Given a dictator of bias B and initial state y, there exists a MPE of the game.*

While there exist multiple equilibria, they differ only in non-substantive changes to $F$'s strategy. Here we will describe the intuition behind the strategies played in this essentially unique equilibrium, which we call $\sigma^*$. A formal description of $\sigma^*$ is in Appendix C.2, as is the proof of Proposition 2.1.

$L$ only accepts bargains today where the value of the bargain plus the discounted value of war tomorrow is at least her war value today. Otherwise $L$ goes to war. Formally, define $H(x,B) = x + \delta M_L(x,B)$ so that the condition for $L$'s acceptance is that $H(x,B) \geq M_L(y,B)$. This is a reasonable policy for $L$ since when $L$ accepts, she does better than or equal to her outside option of war. $L$'s rejection space is based off a pessimistic view of $F$'s future offers: namely, that they lead to a continuation value exactly equal to $L$'s war value. This is, however, the only reasonable assessment of continuation values. This is because $F$'s bargaining power when making take-it-or-leave-it offers leads to a commitment problem: even if $F$ made a promise today to deliver a sequence of offers giving $L$ strictly more than her war value, $L$ has no way of making sure that $F$ carries through on that promise. In fact, if tomorrow $F$ reneged on that promise and offered $L$ only her war value, $L$ still prefers to accept. But, as argued below, $F$ prefers wherever possible to minimize the value that $L$ receives, meaning that $F$ will renege on any promise that gives $L$ a payoff strictly greater than her war value.

Given $L$'s strategy, $F$'s can be broken into two classes of offer: those leading to war immediately and those resulting in $L$'s acceptance. Because $L$'s strategy requires that $F$ promise $L$ at least her minmax value, and because $F$ must guarantee himself at least the value of his outside option, it is possible that $F$ prefers war in the current period. This is the case when $M_F(y) + M_L(y,B) > \frac{1}{1-\delta}$; that is, when the total value demanded by both players exceeds the value of the bargain itself. For clarity, we define the set of all $y$ such that $M_F(y) + M_L(y,B) > \frac{1}{1-\delta}$ as $\mathscr{W}^0(B)$. We call this the *initial war set* and say that in this set the *bargaining space is empty*. If the current state falls in $\mathscr{W}^0(B)$, any offer $F$ may make that garners acceptance by gives $F$ strictly less than his war value, so $F$ prefers to make an unreasonable offer – say $x = 0$ – that leads to war in the current period.

If $y \notin \mathscr{W}^0(B)$, it is possible for $F$ to make an offer that leads to peace. $F$'s problem is to choose $x$ maximizing her value while ensuring that $H(x,B) \geq M_L(y,B)$. Because of the essentially zero sum nature of this game, that value is $x$ such that $H(x,B) = M_L(y,B)$;[45,46]

[45] It is possible that this equation has multiple solutions in $[0,1]$, in which case $F$ choose the minimal solution, retaining for himself as much of the bargain as possible.

[46] It is also possible that for some parameterizations and some values of $y$ and $B$ there is no solution to this equation. When no solution exists to this equation, war results since any sequence of offers by $F$ that $L$ would prefer to war is strictly increasing, thus eventually an offer in this sequence leads to a state where a solution

that is, $F$ makes an offer that guarantees $L$ exactly the value of $L$'s outside option without actually causing war.

A diverse set of outcomes are possible depending on the value of the parameters. The following proposition summarizes these.

**Proposition 2.2.** *For all $\delta$, $C$, and $G$, there exist $y$ and $B$ such that the following are equilibrium outcomes:*

1. *Permanent peace results in a sequence of offers converging to a steady state $y^*(B)$. Furthermore, there is $\overline{B} > 1$ such that for $B \geq \overline{B}$, $y^*(B) > 0$ while for $B < \overline{B}$, $y^*(B) = 0$;*

2. *War results immediately; and*

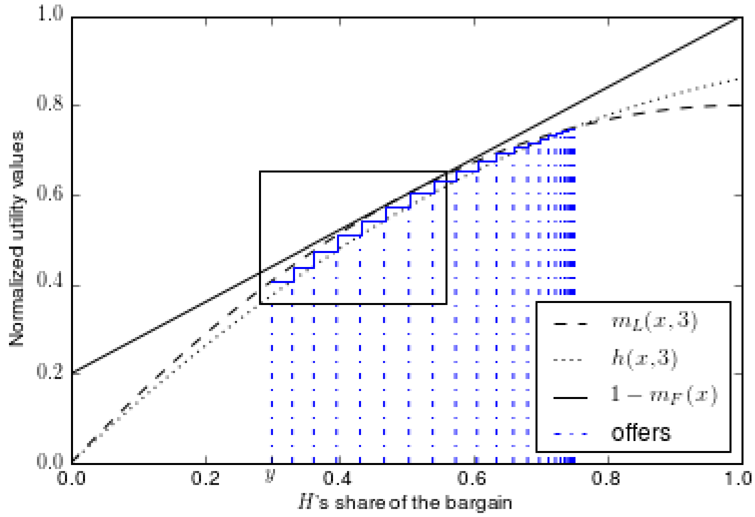
3. *War results in a finite number of periods.*

Figures 1-4 provide illustrations for each of these cases. For clarity, in all figures we normalize utility values by multiplying all payoffs by $1-\delta$. We denote a normalized value function with a lower case version of the same name: for example, $m_L(y,B) = (1-\delta)M_L(y,B)$ and $h(x,B) = (1-\delta)H(x,B)$.

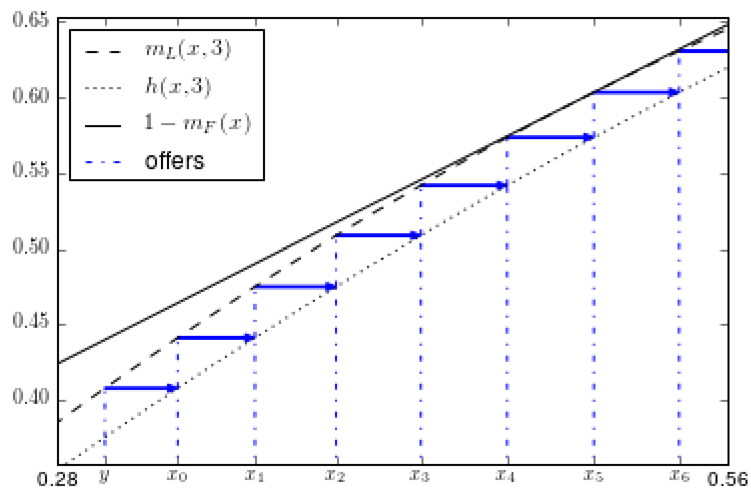
Figures 1 and 2 demonstrate peaceful equilibrium in which offers converge to a steady state. In Figure 1, the leader's bias is sufficiently high that $L$, and thus $H$, eventually captures the majority of the surplus despite an initial distribution favoring $F$. Conversely, Figure 2 depicts an equilibrium in which an unbiased leader gradually gives up more and more of the surplus. Anticipating the results in Section 2.3.2, these two examples show how an unbiased leader – one whose incentives are exactly aligned with those of $H$ – is not the optimal choice of representative.

Figure 3 demonstrates how, for the same initial state $y$ as in Figures 1 and 2, a leader

exists. At this point, $F$ does not have a credible sequence that promises a value greater than $M_L(y,B)$, thus $L$ rejects. Backward induction implies that this sequence is not credible in the current state. See Remark C.1 in Appendix C.2.1 for a more detailed explanation.

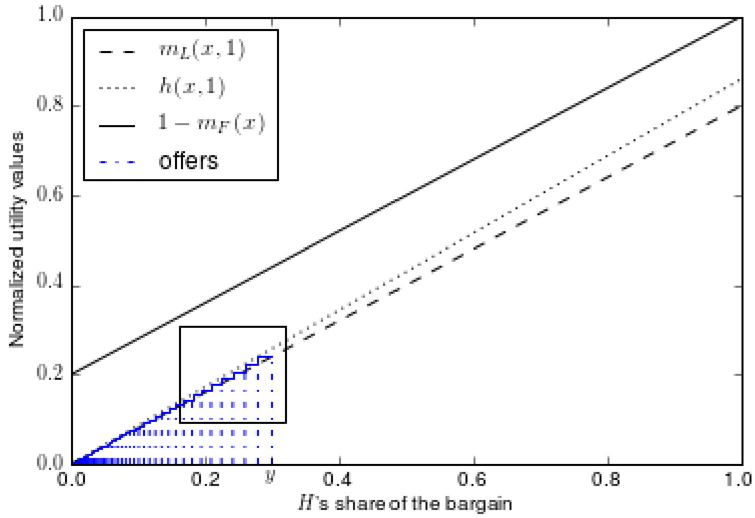



(a) View of the entire bargaining space

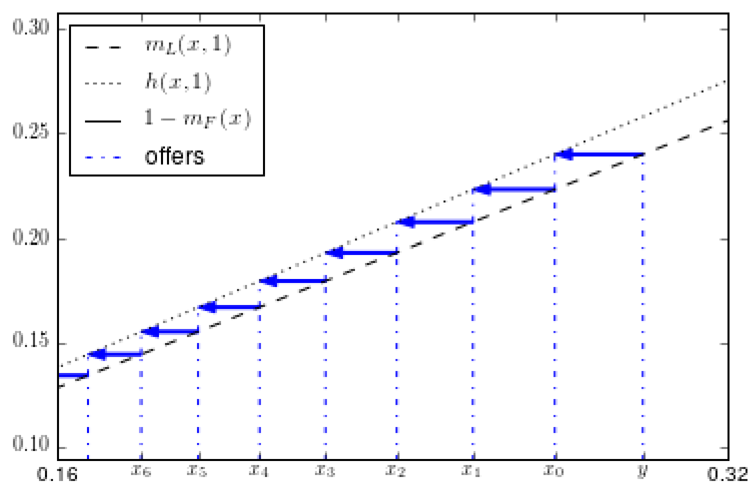



(b) Close up: A graphical representation of how to calculate the first six offers.

Figure 1: A peaceful outcome with a biased leader: $(y,B) = (.1,3)$. Additional parameters: $\delta = .7$, $C = .2$, $G = .4$, and $p(y) = y$. Equilibrium offers must satisfy $M_L(y',B) = x + \delta M_L(x,B)$. An $x$ solving this equality can be found by drawing a horizontal ray from $m_L(y',B)$ (on the dashed arc) toward the curve $h(x,B)$ (the dotted arc) and finding the $x$-coordinate of the point where this ray intersects $h(x,B)$. For instance, in Figure 1b the line segment drawn from $(y, m_L(y,B))$ to $(x_1, h(x_1,B))$ represents the equality $m_L(y,B) = (1-\delta)x_1 + \delta m_L(x_1,B)$.




(a) View of the entire bargaining space.




(b) Close up of the boxed region in Figure 2a.

Figure 2: A peaceful outcome for an unbiased leader: $(y,B) = (.3,1)$. Additional parameters: $\delta = .7$, $C = .2$, , and $p(y) = y$.

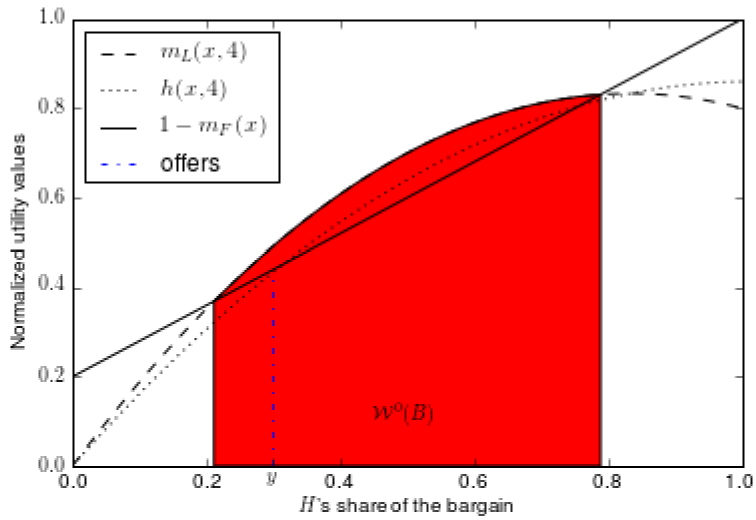



Figure 3: War occurring immediately in equilibrium. where $.3 \in W^0(4)$. Additional parameters: $C = .2$, , and $p(y) = y$.

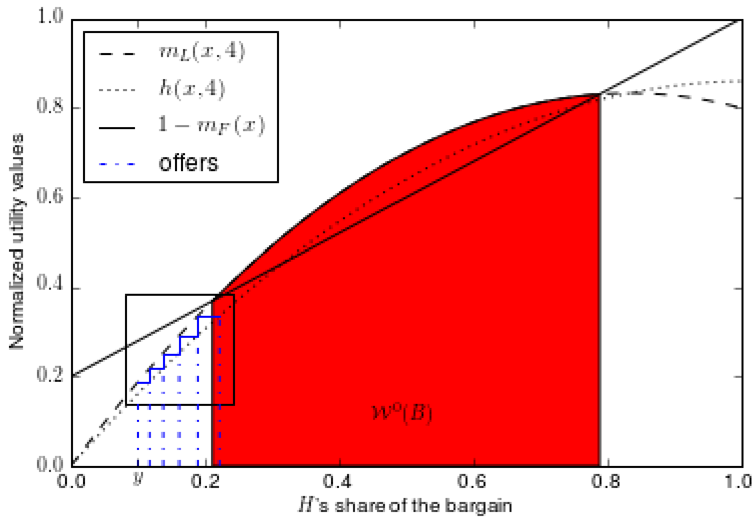
with a too high of a bias leads to war immediately. The shaded set represents the set of initial states where the bargaining space is empty for the given $B$. Finally, Figure 4 shows how, for the same $B$ that leads to war immediately above but for a lower initial state $y$, war still results, though only after a delay of 5 periods.

### 2.3.2 Optimal Bias with Commitment - Perfect Dictatorship

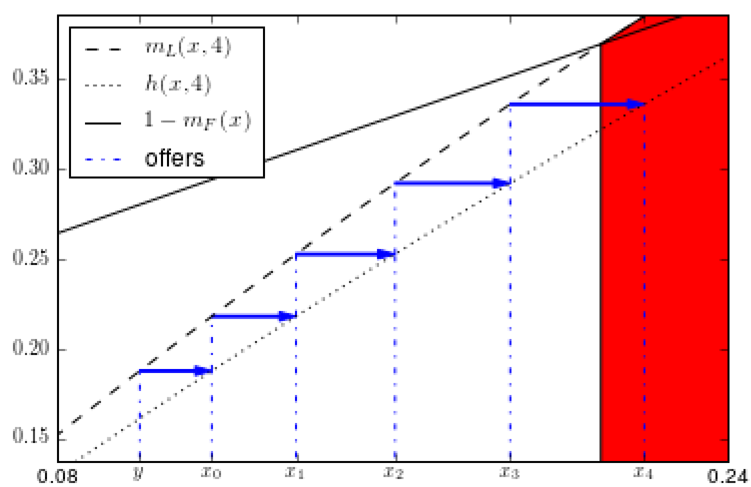
This section analyzes the case where $H$ can commit to choosing a permanent dictator. Let $B^*(y)$ represent the optimal bias for a leader chosen at an initial state $y$. Our first result shows that this optimal leader is always biased towards war.

**Proposition 2.3.** $B^*(y) > 1$ *for all* $y$.

In general, $H$'s choice of the optimal bias equates to the choice of the optimal size of the bargaining space, represented graphically in Figures 1-4 as the distance between $1 - m_F(y)$ and $m_L(y,B)$. Once $B$ has been chosen, the size of the bargaining space is fixed and the equilibrium in Section 2.3.1 has $F$ offering the bargain least favorable to $L$ and $H$.




(a) View of the entire bargaining space.




(b) Close up of the boxed region in Figure 4a.

Figure 4: Delayed war in equilibrium: $(y,B) = (.1,4)$. Additional parameters: $\delta = .7$, $C = .2$, $G = .4$, and $p(y) = y$.

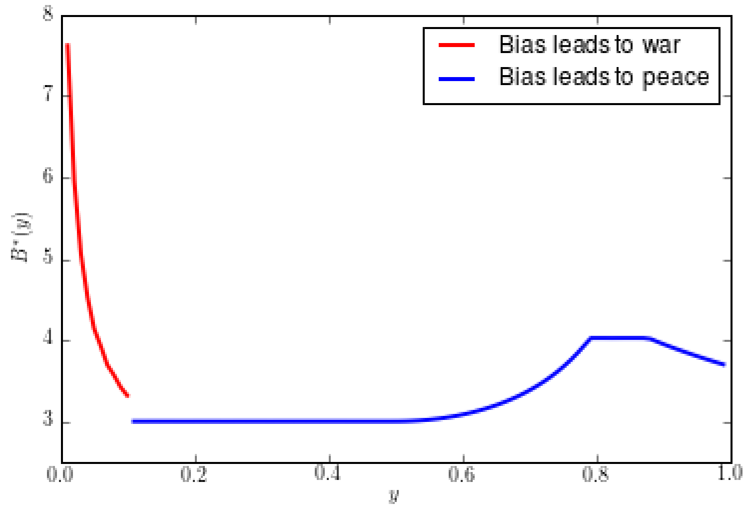



Figure 5: Optimal choice of bias when $\delta = .98$, $C = .2$, $G = .4$, and $p(y) = y$.

However, the size of the bargaining space is decreasing in the bias of the leader. Thus $H$ has an incentive to choose a leader with a high bias. Choice of a bias that is too high, however, leads to war immediately and $H$'s desire to capture at least some of the inefficiency of war leads to an effective cap on which bias it will choose.

A second way of thinking about $H$'s preference for higher biases follows from the fact that, conditional on war not occurring in the current period, higher biases lead to a faster rate of transfer of the bargain to $H$. Because $H$ is impatient, he prefers higher biases today. This preference for higher bias leads to a tension, however, because a bias that is too high leads to war. Furthermore, for a fixed $y$ the delay before war occurs, if war occurs and it does not occur immediately, is decreasing in $B$. $H$, unlike $L$, faces the full inefficiency of war when it occurs and so, all else being equal, prefers that war occur later, if it occurs at all. Thus, there is a tradeoff between acquiring more of the bargain today – the short run benefit of increasing bias – and facing the inefficiency cost of war sooner – the long run cost of increasing bias.

This tension appears in Figure 5, which is typical of such graphical characterizations of

$B^*(y)$. For low $y$, the short run benefit of increased bias outweighs the long run cost: when $H$ has little to begin with, he prefers to rapidly acquire more of the bargain even if doing so leads to war eventually.[47] This is reminiscent of the poverty trap in macroeconomic models. Here, low resource states will be tempted to choose leaders that are destructive to the states long term prospects because this is the exact type of leader that will generate the most improvement in the international bargain initially. In particular, for $y \approx 0$, war, despite its delay, occurs when the the distribution of power is still very much in $F$'s favor.

The tension becomes less extreme as $y$ increases and there is a point (in Figure 5, $y \approx .1$) where the short run benefits due to warlike $B$ no longer outweigh the long run costs. For such $y$ that are less than $1/2$, the biases to which $H$ must limit himself when he desires peace are necessarily very close to $B^\dagger$, the maximal bias such that the bargaining space is always nonempty.[48] For these $y$, $F$ still captures much of the inefficiency of war. When $y \geq 1/2$, $H$ captures at least some of this inefficiency for himself. Prior to the plateau in Figure 5, $H$ actually captures all of the inefficiency of war. For $y$ close to 1 this is not possible. In this region, too high of a bias leads to $F$ acquiring more of the bargain over time and so $H$ again leaves some of the inefficiency of war on the table.

## 2.4 Bias without commitment: Democracy

Our model of democracy is the same as our model of dictatorship, except that the leader can be replaced. How the leader is replaced depends on the domestic institutions within country $H$. An almost unlimited number of models of domestic institutions would fit within our baseline model. We first focus on the optimal democratic outcome from the perspective of $H$. In Section 2.4.1, we show that this leads to peace and to the democracy eventually acquiring all resources within the system. Section 2.4.2 provides a brief comparison of the

[47] It should be noted that in Figure 5, even for low $y$, war never occurs immediately.

[48] For the parameters in Figure 5, $B^\dagger = 3$.

outcomes in optimal democracies and optimal dictatorships. Because the assumptions that lead to the optimal democratic outcome are perhaps untenable, in Section 2.4.3 we provide an example of how a less than perfect democracy leads to war. There, war results when a democracy lacks a commitment technology to not increase the bias of future leaders too quickly. Finally, in Section 2.4.4, we show how political parties can provide a commitment technology that restores peaceful outcomes in a very simple model of domestic politics.

### 2.4.1 Optimal Bias without Commitment - Perfect Democracy

In this section we consider what happens when $H$ can replace the leader in every period. Recall from Section 2.3.2 that $H$'s problem can be thought of finding the bias that minimizes the bargaining space. In particular, conditional on war not occurring, $H$ prefers to give $F$ a payoff as close to $M_F(y)$ as possible and this is achieved by choosing $B$ such that $(1-\delta)M_F(y) = 1-(1-\delta)M_L(y,B)$. However, if $H$ cannot commit to its choice of leader and instead chooses a sequence of biases according to this rule, he gives $F$ much less than his war value.

Without commitment, therefore, $H$'s problem is no longer equivalent to minimizing the bargaining space conditional on balancing the short run benefits and long run costs of high biases. Instead, it involves directly finding a sequence of biases that give $F$ as low a payoff as possible. Indeed, the strategy profile $\sigma^{**}$ below gives $F$ exactly his war value.

**Definition.** Suppose the current state is $y$ and consider a strategy profile $\sigma^{**}$ defined as follows:

1. A leader of bias $B$ accepts any offer $x$ such that $x \geq m_L(y,B)$.

2. Given $(y,B)$, $F$ offers $x = m_L(y,B)$ if

$$1 - m_L(y,B) + \delta M_F(m_L(y,B)) \geq M_F(y)$$

and $x = 0$ otherwise.

3. $H$ chooses $B^{**}(y)$ satisfying

$$B^{**}(y) = \frac{(1-\delta)C + \delta C(1-C)y + G\left[p(y)(1-y) - \delta Cy(1-p(y))\right]}{(1-\delta C)G(1-y)p(y)}$$

for each $y$.

**Proposition 2.4.** *$\sigma^{**}$ is the unique pure strategy Markov equilibrium in the game without commitment. $\sigma^{**}$ is always peaceful.*

That $\sigma^{**}$ is an equilibrium is straightforward. The leader's acceptance rule is the myopic equivalent of the leader's acceptance rule in the commitment equilibrium. $F$'s offer is also as in the commitment equilibrium but with an additional individual rationality requirement which $H$ must satisfy in order to keep $F$ willing to bargain in the future. Finally, $H$'s rule for choosing $B$ keeps $F$ indifferent between going to war and continuing on the equilibrium path. In particular, this implies that $F$'s individual rationality requirement is always satisfied, so war never occurs in equilibrium. It can also be shown that $B^{**}(y)$ determines an offer path such that $H$ accrues more and more over time, with $F$'s share of the resource falling to zero in the limit.

### 2.4.2 Comparison: Stable Dictators and Greedy Democracies

In this section we comment briefly on how equilibrium outcomes in our model differ for dictatorships and democracies. Figure 6 illustrates the main points we wish to make. These outcomes differ drastically depending on whether $H$ is a dictatorship or a democracy. As implied by Proposition 2.4, democratic equilibria are always peaceful and have $H$ always acquiring the entirety of the surplus in the long run. Dictatorial outcomes vary in terms of whether war occurs as well as when and where war occurs, if it occurs, and what the steady

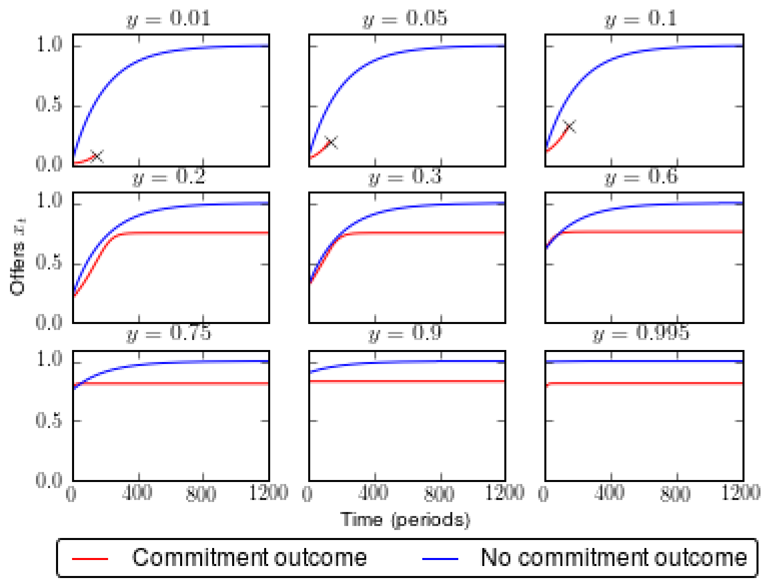



Figure 6: Comparison of equilibrium offers over time for dictatorships and democracies for various initial states $y$. Offer sequences ending with an $\times$ are those which end in war at the given time. Additional parameters: $\delta = .98$, $C = .2$, $G = .4$, and $p(y) = y$.

state is, if it does not.

### 2.4.3 Myopic Democracy

In this section we briefly demonstrate how a democracy in which the electorate does not internalize its commitment problem leads to war immediately. Consider a myopic democracy where the leader is selected so as to maximize the one period bargain received by $H$. $H$ will therefore look for a bias solving

$$m_L(y, B) = 1 - (1 - \delta) M_F(y).$$

Peace would result if $H$ could commit to leader that only demanded the fixed amount $x_t = m_L(y, B)$ in all periods $t$. However, a myopic democracy cannot commit to this. Instead, it will select a new myopically optimal leader every period. Since $F$ remains farsighted, $F$

will make a peaceful offer to such a leader if

$$(1-\delta)M_F(y)+\delta M_F(m_L(y,B)) \geq M_F(y) \Leftrightarrow M_F(m_L(y,B))-M_F(y) \geq 0.$$

Since $M_F(\cdot)$ is decreasing in its argument, this inequality fails to hold if

$$1-(1-\delta)M_F(y) > y$$

If we assume the simplest functional form for $p$, $p(y)=y$, we have that the above inequality holds for all $y<1$.

Therefore war always results in this case. Essentially, a myopic democracy increases the bias of its leaders too quickly not recognizing its own commitment problem. The fact that the leader will be replaced by a new myopically optimal leader is what drives the commitment problem. In general, this type of result can occur in any democracy whose electorate is too short sighted to recognize its own commitment problem.

### 2.4.4 Political Parties as a Commitment Device

In this section we explore how political parties may serve to overcome the commitment problem introduced in last section. If political parties must take a stance on both foreign and domestic issues, the constraints of the commitment problem are relaxed. In particular, if fluctuations in the preferences of the electorate lead to the election different parties with some frequency, some level of aggressiveness can be tolerated in a myopic democracy without causing war.

In a two party system, let party $A$ be myopic in the sense of the previous section. Let the other party, $U$, be unbiased. Every period the parties nominate a new leader representative of their foreign policy preferences. Imagine that the parties are not elected based on their foreign policy choices, but on a domestic issue. Suppose that the electorate favors $A$'s stance on this issue with probability and $U$'s with $1-q$ probability. Maintain the assumption from Section 2.4.3 that $p(y) = y$.

Let $x^i(y)$ be the offer that $F$ makes to when party $i$ is in power. Then, using the fact that $F$ will offer $L$'s minmax value, it can be shown that

$$x^A(y) = (1-C)y + C \text{ and } x^U(y) = (1-C)y.$$

Note that $A$ captures all of the inefficiency of war, while $U$ captures none. The value functions for $F$ when $A$ and $U$ are in power are, respectively,

$$V_F(A,y) = 1 - x^A(y) + \delta\left[qV_F\left(A, x^A(y)\right) + (1-q)V_F\left(U, x^A(y)\right)\right]$$
$$V_F(U,y) = 1 - x^U(y) + \delta\left[qV_F\left(A, x^U(y)\right) + (1-q)V_F\left(U, x^U(y)\right)\right].$$

Note that as $q \to 1$, $F$'s payoffs approach those discussed in the previous section.

**Proposition 2.5.** *There exist $\overline{q}$ and $\underline{q}$ such that if $q \geq \overline{q}$, war results immediately in all states $y$. If $q \leq \underline{q}$ peace holds for all states $y$.*

*Proof.* The proposition then follows from $q = 1$ case, discussed in Section 2.4.3, and the $q = 0$ case, which is easily shown when $p(y) = y$. □

One could interpret $q$ as resulting from the institutional structure of $H$'s political system, namely from the degree to which the leader is insulated from the electoral implications of domestic politics. The greater the degree of insulation, the more frequently the leader myopically optimal foreign policy party is elected. The optimal peaceful institutional structure is then found by solving for $q^* = \max\{\underline{q}\}$.

The result that political parties can operate as a commitment device within a democracy elucidates the complexities of commitment within this framework. Dictatorships – even optimal ones – go to war precisely because society is able to commit to the dictator and lacks the flexibility of replacing leaders that democracy affords. One may view a democracy where a perfectly forward looking electorate optimizes their leader only over the nature of the leader's international bargaining preferences as impractical. Yet, as Section 2.4.3 has shown, introducing myopia in the selection of leader's can make the problem of war even more severe than in the case of dictatorship. Political parties allow a democracy to institutionalize a commitment to selecting less belligerent leaders with a certain probability while avoiding the perils of making a permanent commitment to a dictator. This result relies on an electorate that is at least somewhat myopic and inward looking. If this is indeed the case, party competition serves as a commitment device preventing too much flexibility in leader selection that also allows a certain degree of leader replacement, thus avoiding over commitment to a particular bargaining path.

## 2.5 Conclusion

It is unsurprising that with a high enough bias in favor of war, a dictator may lead a country into war even when bargaining is possible. However, it is surprising that a dictator who is optimally chosen by a country's electorate would also have this property. Interestingly, such a war may not occur immediately as an outside country may find it optimal to appease such a leader for a period of time. On the other hand, there exists a credible peaceful path for a democratic country to capture all of the inefficiency of war. The result requires that leaders be selected by an infinitely forward looking populace that is able to optimize their leader on the basis of her foreign policy alone. The strictness of this possibility result demonstrates the ease with which myopic behavior on the part of the populace leads to war even in democracies. Party competition over a platform of domestic and foreign policy can serve as an institutional commitment device within a democracy which leads to a peaceful outcome so long as the less biased party is selected with enough frequency.

Throughout this chapter, we have assumed that there is no uncertainty other than over the outcome of a war in any given state. Our model could easily be extended to allow for imperfectly observed leader biases. Here we note only that unbiased leaders of $H$ have an incentive to pretend to have a higher bias[49] and that highly biased leaders may prefer to downplay their bias if doing so leads to their being in a better position to wage war in the future. A thorough analysis of incomplete information scenarios in our model potentially has a great deal of explanatory power.

We have also to this point assumed that $F$ had no choice regarding its leader. Extending our model to allow $F$ such a choice may provide valuable insight; in particular, it would allow commentary directly on the phenomenon of dyadic democratic peace. If leaders are selected simultaneously, there would be a positive probability of war as there would exist a risk-return trade-off from the selection of biased leaders. Sequentially elected leaders would introduce first mover advantages, but presumably reduce the risk of war. Party competition may still serve as a commitment device for peace within democracies since more aggressive leaders will credibly be replaced by easier negotiators in the future. Finally, combining two-sided selection of leaders with incomplete information over leader type has the potential to produce interesting results on how posturing and punishment in the bargaining process interacts with the replacement of leaders.

[49]This is true for the same reasons that $H$ itself does not select an unbiased leader in the game with commitment.

# 3 Multilateral War Bargaining

## 3.1 Introduction

Fearon (1995) argues that in order to explain why costly wars occur between rational, unitary states, one must explain why an efficient bargain could not be found. There must be a friction in the form of incomplete information or a commitment problem on the part of at least one of the states. Powell (1999, 2004, 2006) further explores the commitment problem in several settings and identifies a common mechanism – large and rapid shifts in the distribution of power. Fearon and Powell argue that war is caused when one state is expected to experience a large increase in power in the near future. Since no world government exists to enforce a contract, the rising power cannot commit to give the declining power its current war value in the future.

This is a powerful insight into the causes of war, yet the bargaining model of war is typically explored with only two states or in a highly structured setting with more than two states where each state has a role (for instance: attacker, defender, potential ally). Moreover, models involving more than two players tend to not account for dynamics in the system. Without allowing for more than two states, the bargaining model of war cannot hope to explain multilateral interactions like alliance formation. Without allowing for dynamics, multiplayer models miss critical caveats to the static results.

Furthermore, there are many historical cases of war that are rather difficult to describe without discussing multilateral interactions. In the two-player model, a rising state, all else being equal, does not have an incentive to initiate war. Yet major historical cases defy this logic. The rising Prussian state in the 1860s and the rising German state in the 1930s seems to have been the initiators of the Franco-Prussian war and World War II respectively. Furthermore, it's difficult to follow the logic of the Nazi invasion of Russia in 1941 or the Japanese attack on China in 1894 without reference to the impending invovlement of the

U.S. in the European system or of Russian in the East Asian system.

This chapter builds a complete information, three-state bargaining model of war as an initial step towards the construction of a truly multi-state model. While the model in this chapter is only a first step towards building a model of the complexity necessary to handle historical cases, the leap from two players to three is a critical progression. In a two-player model, the effects of alliances and alliance shifting is simply lost. With a model of three players, one can ask with parsimony most of the questions that an *n*-player model would more fully address.

Using this model I make three contributions. First, I describe a method for predicting whether balancing or bandwagoning alliances will form in a static, three-player bargaining model of war. Second, I show two new mechanisms for war. The first mechanism shows that war can occur when a third player is anticipated to enter a two-player system. The initial two players may go to war in order to prevent an adverse alliance forming and to acquire resources in bargaining with the entering player. In the second mechanism, unanticipated changes in the distribution of power can lead to war in a three-player game through alliance switching. In some cases, alliance shifting leads to war under quantitatively different circumstances than in the two-player model. However, alliance shifting can lead a state to initiate war even when rising in relative power, which is qualitatively different from the two-player model. Third, I show that anticipated shifts in the distribution of power affect whether a bandwagoning or balancing alliance forms.

This last contribution makes clear that alliance formation is affected by the anticipation of shifts in resources and not just the current distribution of resources. This is the case when the alliance that would occur in a static world in the given period will shift in the future causing war. If states anticipate this shift before alliances are formed, then states may not form the static alliance in order to avoid a war outcome. The implication is that a peaceful equilibrium requires states to balance the current distribution of capabilities, but also the dynamics of the system. Interestingly, a dynamically balanced system may look like bandwagoning in the short term. Hence, states may balance with a declining hegemon as opposed to a rising weak power in order to avoid war caused by an anticipated alliance shift that would occur once the hegemon is replaced. For instance, the lack of hard balancing against the U.S. becomes far less mysterious when the dynamics of a rising China are considered.

To my knowledge, the modeling technique used here is unique in the formal war literature. However, it has many antecedents. The literature on coalition formation and alliance formation is very large. Two recent surveys in particular are particularly helpful in unifying this vast literature. Ray and Vohra (2013) focuses on the more abstract setting of coalition formation whereas Block (2012) explores the application to alliance formation in particular. However, it is important to distinguish the model in this chapter from several similar papers in alliance formation. Niou and Ordeshook (1986) take an axiomatic approach to alliance formation. Their model predates the bargaining model of war and focuses on rather different notions of stability than considered here. For instance, one notion, "system stability," is just the notion that war would not occur in a three-player environment without incomplete information or commitment problems. This result also holds in this setting. Wagner (1986) provides a non-cooperative model similar in direction to Niou and Ordeshook (1986). Skaperdas (1998), Noh (2002), and Garfinkel (2004) examine interesting non-cooperative, three-player models of alliance formation. However, these models focus more on sharing rules within alliances and lack bargaining and resource dynamics in the sense of this chapter.

Wagner (2007) investigates a three-state bargaining model of war. The focus is on incomplete information as opposed to commitment problems caused by resource dynamics and shifting alliances. However, Wagner makes several points that are important for a complete information model. Of particular importance, one must be aware of what happens after an initial war in a three-player game. Namely, if a state allies and goes to war with the stronger of the two other powers, then, if victorious, that state will be paired in a two-player game with a stronger adversary. Thus, the short run benefits of a stronger ally are at least somewhat countered by a weaker bargaining or military situation in the future. In general, a model must chose to the degree in which allies who win a war can contract over dividing the spoils. On one extreme, the division is entirely determined endogenously by the resources of the two states. In the other, states can contract to divide the resources in any way. For instance, the division of Europe at the end of WWII was partially determined by power realities – Poland falling fully into the dominion of the Soviet Union – and partially by contract – the division of Germany in separate zones. This chapter's model addresses this issue in a reduced form manner. Alliances are assumed to have a certain value for each member but can vary in any way that corresponds to some regularity conditions. Thus, most assumptions on the future division of resources can be fit into this formulation.

The next section constructs the static model. This model is then characterized in terms of which alliances will form depending on the distribution of resources, functional forms, bargaining protocols, and sharing rules. Next, the model is made dynamic and various implications are explored. First, I look at the expected entrance of third player into an initially two-player system. Then, unanticipated resource shocks in a three-player system are explored. Finally, I examine the impact of anticipated shocks on alliance formation.

## 3.2 Static Model

In this section I specify a cooperative game, labeled $\Gamma$. There are three players labeled $a$, $b$, and $c$.[50] Each player is endowed with a single resource, $r_i$, with the magnitude ordering $r_a \geq r_b \geq r_c$. $r$ represents the vector $\{r_a, r_b, r_c\}$. Total system resources are represented by

[50] When discussing the general model, I will insert "player" for "state" in order to emphasize the general applicability of the model and to avoid confusion with state variables.

$R = r_a + r_b + r_c$. As is commonly done in the bargaining literature, resources are assumed to be an inherent characteristic of each player that is not freely transferable. However, the utility derived from resources is freely transferable.[51] An example of this would be the resource of California in the U.S. California provides the U.S. federal government with a large source of tax income each year. That income can be spent on the military, consumed within the U.S. on domestic programs, or transferred as aid to other countries. Hence, the resource of California simultaneously provides the U.S. with power and utility – utility which is transferable to other countries. On the other hand, California itself and not just the stream of utility from California may be viewed as transferable, but the political costs of transferring California to another country are extremely high. Here, I simply treat those costs as infinite and that resources themselves can only be acquired through war.

The alliance structure of a system is represented by the state variable $s$. There are 4 possible states: $s_0$ where there are no alliances; $s_{ab}$, $s_{ac}$, and $s_{bc}$ where the two labeled players are in an alliance together. $S$ is the set of all these state variables. I will refer to an alliance that includes the strongest state $a$ as a bandwagoning alliance and an alliance between the two weaker states, $b$ and $c$, as a balancing alliance. Alliances are marked by two principal characteristics: (1) alliance members make a (one period for the dynamic version) binding commitment to go to war on the side of its partner regardless of who initiates war; (2) alliance members can commit to intra-alliance transfers of utility in war and peace. The first characteristic may be justified as assuming that signing an alliance contract involves sufficient audience costs to fully commit a player to an alliance for a period. The second characteristic represents the assumption that players are free to divide a bargain over a defeated player's resources according to a prearranged contract for the static period. However, the flexibility of this contract is limited by the bargaining power within the alliance. The model assumes that any contract is possible, but endogenously constrains

[51] See Chadefaux (2011) for a critique of this notion.

the contract so that it properly reflects each player's outside option to the alliance. This outside option is quite complex since it is dependent on the values of all other alliances which are themselves endogenous to each other.

In state $s_0$ where no players are allied, utilities are represented by the vector $v(s_0) = \{v_a(s_0) \geq 0, v_b(s_0) \geq 0, v_c(s_0) \geq 0\}$ where $\sum_i v_i(s_0) \leq R$. This setup for $s_0$ is agnostic as to whether this vector of utilities is generated by peaceful transfers or war. I have not allowed for the possibility of a grand coalition since no player can threaten another in order to receive a transfer in such a state. However, $s_0$ could be seen as representing this type of structure in some contexts.

Before specifying utilities in a state $s \neq s_0$, one first needs to introduce the war values in these states. The alliance and outside player have war values designated $w_{ij}(s_{ij}) \geq 0$ and $w_z(s_{ij}) \geq 0$ respectively. War values are the payoff if any player initiates war. Hence, they serve as an outside option for each state $s$. A particular specification of this model could determine war values using many possible functional forms that fit the assumptions of the model made below. Throughout, I will analyze a leading example where war values are determined by a costly lottery as is common in the literature.

I assume that $w_{ij}(s_{ij}) + w_z(s_{ij}) < R$ so that war is inefficient. Let $\gamma(s_{ij}) = R - w_{ij}(s_{ij}) - w_z(s_{ij})$ represent the inefficiency cost of war. Note that the inefficiency of war depends on the state of alliances. For instance, a war of overwhelming force may be less inefficient than one between equals. War is assumed to be game ending. Unlike transfers, which are just in utility, war results in the winner acquiring all the actual resources of the defeated player. This distinction becomes important when looking at the dynamic model. Note also the simplification that I am not allowing war to result in the transfer of only a portion of the resources of the loser to the winner.

I assume that there exists a known bargaining protocol $\rho(r,s) \in [0, \gamma(s)]$ that determines the outside player's bargaining power and hence how much of the inefficiency of war the outside player captures in bargaining. This is a quite general way to handle bargaining. For instance, a bargaining protocol where the alliance makes a take-it-or-leave-it offer would imply that $\rho(r,s) = 0$. In the other extreme, the outside player makes the take-it-or-leave-it offer and $\rho(r,s) = \gamma(s)$. Bargaining protocols where $\rho(r,s) \in (0,\gamma(s))$ represents situations where neither the alliance or outside player have all of the bargaining power. Note that the range of possible bargains is determined by the level of inefficiency, which only depends on the state $s$. However, the bargaining protocol can vary generally with $r$ in combination with $s$. Hence, the bargaining power of an alliance can differ for two reasons. First, different levels of resources shift the alliance's war value and hence the minimum level achievable in a bargain. Second, different alliances, with different amounts of resources may have different levels of bargaining power within the acceptable range of bargains.

For any particular specification, I assign values of a player being in an alliance or being an outside power. Since individual players face an externality when others form an alliance, $\Gamma$ must technically be specified in partition function form. When the alliance has value

$$u_{ij}(s_{ij},\rho) = w_{ij}(s_{ij}) + \big(\gamma(s_{ij}) - \rho(r,s_{ij})\big)$$

the outside player has value

$$u_z(s_{ij},\rho) = w_z(s_{ij}) + \rho(r,s_{ij})$$

and $v(s_0)$ represents values in the state without alliances. Note that these transfers arise from the threat of war. However, in war $u_{ij}(s_{ij},\rho) = w_{ij}(s_{ij})$ and $u_z(s_{ij},\rho) = w_z(s_{ij})$, so that war is inefficient. Therefore, war does not occur in the static, cooperative model and one can safely use the non-war utility specifications.

Assumption 1 puts some natural structure on war and alliance values.

**Assumption 1.** *The following regularity conditions are satisfied:*

*1.) If $r_i > r_j$, then $w_i(s) > w_j(s)$.*

*2.) $w_{ij}(s_{ij}) > v_i(s_0) + v_j(s_0)$.*

*3.) $v_i(s_0) > u_i(s_{jz}, \rho)$ for all $\rho$.*

A1.1 implies that war values are increasing in resources. This is assumed to hold for both individuals and alliances. The assumption is natural since resources represent a player's war making power. A1.2 implies that the total war value of an alliance between two players is higher than the value to those players alone in $s_0$. This implies that there is at least a marginal gain in total for coordinating actions in an alliance. A1.3 implies that a player always does worse when the other players coordinate then when they are independent.

At this point it is useful to fix ideas with an example.

**Example.** (Leading Example) First, I will specify war values. If war starts, the probability of victory by an alliance $ij$ is $q_{ij} = \frac{r_i + r_j}{R}$ or for outside player $i$ is $q_i = \frac{r_i}{R}$. Cost to warring players in $s \neq s_0$ is $\gamma(s_{ij}) = \frac{r_c}{2\max\{p_z, 1-p_z\}}$. Thus, war costs are bounded above by $c$'s resources and decreasing in the degree of the preponderance of resources possessed by the stronger power. This means that the war values for an alliance and outside player are

$$w_{ij}(s_{ij}) = q_{ij}\left(r_i + r_j + r_z\right) - \frac{r_c}{2\max\{p_z, 1-p_z\}} = r_i + r_j - \frac{r_c}{2\max\{p_z, 1-p_z\}}$$

and

$$w_z(s_{ij}) = q_z\left(r_i + r_j + r_z\right) - \frac{r_c}{2\max\{p_z, 1-p_z\}} = r_z - \frac{r_c}{2\max\{p_z, 1-p_z\}}$$

respectively. I assume that $\rho(r,s) = 0$, thus alliances always capture the bargaining surplus.[52] Utilities are then

$$u_z(s_{ij}) = w_z(s_{ij}) = r_z - \frac{r_c}{2\max\{q_z, 1-q_z\}}$$
$$u_{ij}(s_{ij}) = r_i + r_j + \frac{r_c}{2\max\{q_z, 1-q_z\}}$$

I assume that $v_i(s_0) = r_i - \gamma(s_{jz}) + q_i \sum_{s \neq s_0} \gamma(s)$. One can motivate this by thinking that in $s_0$ a war would result in each player paying their outside player cost and winning the war with probability $q_i$. However, players avoid war through a bargaining mechanism in which they have bargaining power proportional to their probability of winning a war.

This leading example is just one of many possibilities, but is convenient to work with. It is easy to verify that it fits the assumptions I have made so far. It is true that $w_{ij}(s_{ij}) + w_z(s_{ij}) < R$ since players pay a cost when going to war. A1.1 holds since $w_{ij}(s_{ij})$ is increasing in both $r_i$ and $r_j$, while $w_z(s_{ij})$ is increasing in $r_z$. A1.3 holds since

$$v_i(s_0) = r_i - \gamma(s_{jz}) + q_i \sum_{s \neq s_0} \gamma(s) > r_i - \gamma(s_{jz}) = u_i(s_{jz})$$

Since, A1.3 holds and every state is efficient here (no war), A1.2 also holds since $u_{ij}(s_{ij}) = R - u_z(s_{ij}) > R - v_z(s_0) = v_i(s_0) + v_j(s_0)$.

The final element of the static model is specifying the constraints on the internal sharing rules of an alliance. First, individual utility within an alliance is constrained by the total utility the alliance produces so that $u_i(s_{ij}, \rho) + u_j(s_{ij}, \rho) = u_{ij}(s_{ij}, \rho)$. However, this restriction alone does not ensure that every sharing rule properly accounts for the outside options available to a potential alliance member. Fortunately, it turns out that it is easy to modify any intra-alliance utility sharing arrangement so that it is consistent with each

[52] This is arbitrary, but could be thought of as a sort of voting rule.

player's endogenous outside option. This is done by allowing players within alliances to unilaterally modify their share of the alliance utility downward.

When in a given state $s \neq s_0$, an allied player, $i$, can commit to offering $x_i$ resources to the other allied player, $j$. Let final allied utilities be specified as

$$v_i(s_{ij}, \rho) = u_i(s_{ij}, \rho) - x_i$$
$$v_j(s_{ij}, \rho) = u_j(s_{ij}, \rho) + x_i$$

Note that it is still the case that $v_i(s_{ij}, \rho) + v_j(s_{ij}, \rho) = u_{ij}(s_{ij}, \rho)$.

Allowing a player to unilaterally commit to transferring resources in a given state is a natural modification. The game $\Gamma$ allows the modeler to specify any consistent sharing rule. This modification simply ensures that the sharing rule is in fact consistent with each player's outside option. If both allied players in state $s$ value that state the highest, then there is no need for a transfer. However, one player, $i$, may value that state extremely highly while the other player, $j$, prefers another state, $s'$. If $i$ can unilaterally transfer utility to $j$ such that $j$ prefers to stay in state $s$, then $i$ has every incentive to do so. The original utility specification simply did not account for these incentives properly.

However, one might wonder why players are allowed to commit to a transfer in the current state, but not be able to promise a transfer if the player switches alliances. This reflects the notion of commitment in this game. A player is allowed to unilaterally change upwards the amount they transfer to another player in a given state, but not allowed to commit to a promise to do this in "the future" after a switch in alliances. In fact, the player who made the promise would have no incentive to keep it, if it did not correspond to type of transfer that is already allowed here. Essentially, this specification takes the view that when considering a state $s$, all other states $s'$ offer known utility values. Only $s$ can be credibly modified.

To see how to construct sharing rules in practice, the following example introduces a sharing rule to the leading example.

**Example 3.1.** Let the value to an individual within an alliance be the expected value of the following game.[53] If the state is $s_{ij}$, then with probability $p(r_i, r_j)$, $i$ proposes a sharing rule $\phi_i(s_{ij}) \in [0,1]$ where $i$ receives utility $u_i(s_{ij},\rho) = \phi_i(s_{ij})u_{ij}(s_{ij})$. $j$ plays "accept" or "reject." Alternatively, with probability $1 - p(r_i, r_j)$, $j$ proposes a sharing rule $\phi_j(s_{ij}) \in [0,1]$ where $i$ receives utility $u_i(s_{ij},\rho) = \phi_j(s_{ij})u_{ij}(s_{ij})$. $i$ plays "accept" or "reject." Rejection implies the alliance breaks causing $s_0$ to form. Assume that $p(r_i, r_j)$ is weakly increasing the first term and weakly decreasing in the second.

Let $E_0$ be the expectation operator over who gets to propose before an alliance forms. Applying this assumption, through backwards induction the value of being in an alliance can be calculated according to Lemma 3.1.

**Lemma 3.1.** *In Example 1,* $E_0 u_i(s_{ij},\rho) = p(r_i,r_j)\left(u_{ij}(s_{ij}) - v_j(s_0)\right) + \left[1 - p(r_i,r_j)\right] v_i(s_0)$ *for all* $i$, $j$.

Lemma 3.1 leads immediately to the following lemma:

**Lemma 3.2.** *In Example 1,* $E_0 u_i(s_{ij},\rho) > u_i(s_{jz},\rho)$ *for all* $\rho$ *and for all* $i$, $j$.

Lemma 3.2 shows that states always prefer to be in an alliance to being outside an alliance in this setup.

I can now write the expected values of the various alliances to the players. I refer to this type of table as an alliance-utility table.

[53]This processes produces a value for an alliance that is akin to a Shapley Value modified to give player's proposer power that is dependent on their relative resources.

Table 3: Expected Payoffs

| | $s_0$ | $s_{ab}$ | | $s_{bc}$ |
|---|---|---|---|---|
| $u_a$ | $v_a(s_0)$ | $E_0u_a(s_{ab},\rho)$ | $E_0u_a(s_{ac},\rho)$ | $w_a(s_{bc})+\rho(r,s_{bc})$ |
| $u_b$ | $v_b(s_0)$ | $E_0u_b(s_{ab},\rho)$ | $w_b(s_{ac})+\rho(r,s_{ac})$ | $E_0u_b(s_{bc},\rho)$ |
| $u_c$ | $v_c(s_0)$ | $w_c(s_{ab})+\rho(r,s_{ab})$ | $E_0u_c(s_{ac},\rho)$ | $E_0u_c(s_{bc},\rho)$ |

This can be made more concrete by assuming that proposer power within an alliance is proportional to a player's size within an alliance. Thus $p_i(r,s_{ij}) = \frac{r_i}{r_i+r_j}$. After some algebra, one can then calculate individual utilities within an alliance as

$$u_i(s_{ij}) = \frac{r_i}{r_i+r_j}\left[u_{ij}(s_{ij}) - v_j(s_0)\right] + \frac{r_j}{r_i+r_j}v_i(s_0)$$

Given these functional forms, it is easy to calculate the values as in Table 3 for any particular distribution of resources by repeatedly plugging in. Table 4 gives an example of rounded payoffs for a particular vector.

Table 4: Expected Payoffs for the resource distribution: $r_a = 50$, $r_b = 30$, $r_c = 20$

| | $s_0$ | $s_{ab}$ | $s_{ac}$ | $s_{bc}$ |
|---|---|---|---|---|
| $u_a$ | 52 | 55 | 57 | 40 |
| $u_b$ | 30 | 32 | 23 | 37 |
| $u_c$ | 18 | 14 | 20 | 23 |

In this case, both *b* and *c* prefer an alliance with each other rather than with *a*. Thus, it is natural to predict that a balancing alliance $s_{bc}$ forms. Balancing alliances are relatively attractive here because alliances have proposer power in inter-alliance bargaining. Due to this, the benefit of having to share less of the alliance resources when allied with a smaller partner outweighs the decrease in overall alliance value. For instance, while the total alliance value of $s_{ab}$ in Table 4 is 87, *b* prefers to be in an alliance *c* which only has a total alliance value of 60.

In the above example it seemed natural to predict that the balancing alliance *bc* would form. More generally, the static model should predict alliance formation for any coherent specification of player utilities in the alliance-utility table. The model makes predictions according to a reasonable notion of the stability of an alliance. Define stability in this setting as follows:

**Definition 3.1.** $s \neq s_0$ is stable if there is not a strictly positive deviation for any allied player and the outside player to another state $s'$. $s_0$ is stable if no two players have a weakly positive deviation to an alliance state.

This says that $s_0$ is only stable if no two states prefer to form an alliance and exploit the third state. Now, consider a state $s \neq s_0$. First consider deviations to war. Since war is inefficient, so long as the bargain is within the range specified by $\rho(\cdot)$, war is not a profitable deviation. Therefore, for any bargaining protocol, $\rho(\cdot)$, both the alliance and outside player prefers state $s$ and bargaining protocol $\rho(\cdot)$ to war. All bargaining deviations within the peaceful range are particular to the specified bargaining protocol, however, $\rho(\cdot)$ captures all peaceful possible bargaining outcomes.

Definition 3.1 appropriately handles deviations from $s$ to $s'$ because the outside player always prefers $s_0$ to being isolated. Furthermore, a new alliance state can only form if both parties to the alliance prefer it to the current state. The outside player is not allowed to deviate to another state on her own since she can not force the other players to break their alliance. Allied players only prefer to break the alliance if there is another state that is strictly preferred to the current state.

For example, in Table 4 of Example 1, $s_{bc}$ is the uniquely stable. $s_0$ is not stable since $a$ and $b$ prefer to deviate to $s_{ab}$. $s_{ab}$ is not stable since $a$ and $c$ prefer to deviate to $s_{ac}$. $s_{ac}$ is not stable since $b$ and $c$ prefer to deviate to $s_{bc}$. For $s_{bc}$, $a$ prefers to deviate to any other state, but neither alliance member has a positive deviation to any other state. Hence, $s_{bc}$ is uniquely stable.

The next section characterizes the general implications of stability in predicting balancing and bandwagoning behavior.

## 3.3 Characterization of the static model

The leading example used a particular process to generate Table 4. However, it is useful to work with the alliance-utility table values directly. The only constraint on these values is that they satisfy the regularity assumptions of the model. With this in mind, consider the specifications in Table 4 where the internal sharing rules in states $s_{ab}$ and $s_{bc}$ are changed so the alliance-utility table becomes

Table 5: Modified sharing in $s_{ab}$ and $s_{bc}$

| | $s_0$ | $s_{ab}$ | $s_{ac}$ | $s_{bc}$ |
|---|---|---|---|---|
| $u_a$ | 52 | 53 | 57 | 40 |
| $u_b$ | 30 | 33 | 23 | 32 |
| $u_c$ | 18 | 14 | 20 | 28 |

Without internal transfers, it is clear that no stable state exists in Table 5. *ab* has a positive deviation from $s_0$ to $s_{ab}$. *a* has positive deviation from $s_{ab}$ to $s_{ac}$. *c* has a positive deviation from $s_{ac}$ to $s_{bc}$. Finally, *b* has a positive deviation from $s_{bc}$ to $s_{ab}$. This result is unsurprising. Essentially it reiterates that a core might not exist with three players in this type of setup. Note, to say the core is empty is to make no prediction as to the outcome in this setup. It does not imply war. War would only be implied if it were itself a stable outcome in some state *s*, which is never the case here.

Then which state would be predicted in Table 5 using the committed transfers $x$? *b* would need to offer *a* $x \geq 4$ to remain in $s_{ab}$, but then *b* would prefer $s_{bc}$. *a* could offer *c* $x \geq 8$ not to switch to $s_{bc}$, but then *a* would prefer $s_{ab}$ without transfers. In $s_{bc}$, if *c* offers *b* $x > 1$, *b* prefers to stay in $s_{bc}$ rather than switching to $s_{ab}$ and *c* prefers this transfer to switching to $s_{ac}$. Hence, the cooperative model would predict $s_{bc}$ as the stable state. The final alliance-utility table (with integer values for convenience) would become

Table 6: Final $v$ values for Table 5

| | $s_0$ | $s_{ab}$ | $s_{ac}$ | $s_{bc}$ |
|---|---|---|---|---|
| $v_a$ | 52 | 53 | 57 | 40 |
| $v_b$ | 30 | 33 | 23 | 34 |
| $v_c$ | 18 | 14 | 20 | 26 |

Essentially, this is saying that the modification of allowing committed transfers in any given state produces a stable one in this case. Proposition 3.1 shows that a stable state exists in all cases conforming to A1.1 through A1.3.

**Proposition 3.1.** *Under Assumption 1, every cooperative game* $\Gamma$ *has a stable state* $s \in S$.

It should be noted that the stable state in $\Gamma$ need not be unique. Namely, this is possible when two states give a player the same amount of utility. In these cases, the method here simply predicts a set of stable states. However, the result is still quite striking. With a very reasonable restriction that alliance sharing rules appropriately accommodate the outside option of the allied players, the problem of an empty core disappears and a reasonable notion of stability can be used to predict alliance behavior for any model that fits the relatively weak assumptions of this model.

The rest of this section shows that under an additional assumption, balancing and bandwagoning can be predicted in two large classes of models. The necessary assumption is as follows:

**Assumption 2.** *($s_0$ compatibility)* $v_i(s_{ij},\rho) > v_i(s_0)$ *and* $v_j(s_{ij},\rho) > v_j(s_0)$ *for all* $\rho$.

Assumption A1.2 assures that this assumption is always mutually compatible with $v_i(s_{ij},\rho) + v_j(s_{ij},\rho) = u_{ij}(s_{ij},\rho)$. A2 requires that players within an alliance both prefer the alliance to $s_0$. If A2 does not hold, then $s_{ij}$ is not stable and in isolation from the other possible alliances, the players within alliance $ij$ would have an incentive to alter the sharing rule within the alliance to make it stable to deviations to $s_0$. This is always possible by assumption A1.2. The practical effect of A2 is to rule out $s_0$ as a stable state. This is not always a reasonable restriction, but is efficacious here since I am interested in predicting balancing and bandwagoning behavior.

**Proposition 3.2.** *Given assumptions A1 and A2 and if* $u_i(s_{ij},\rho) > u_i(s_{iz},\rho)$ *and* $u_j(s_{ij},\rho) > u_j(s_{jz},\rho)$*, then* $s_{ij}$ *is uniquely predicted in* $\Gamma$.

This is the straightforward logic used in predicting that Table 4 would imply the balancing state $s_{bc}$. Notice that it is appropriate to use $u_i(\cdot)$ for the value within an alliance here since no transfers, $x_i$, would be made here since both alliance members strictly prefer the alliance already.

In order to find when Proposition 3.2 applies, it will be useful to define the following two scenarios leading to different alliance formation in $\Gamma$:

$\mathbb{U}$: $u_i(s_{ij},\rho)$ is strictly increasing in $r_j$.

$\mathbb{B}$: $u_i(s_{ij},\rho)$ is strictly decreasing in $r_j$.

Allying with a larger player creates more total resources for an alliance since the alliance is able to threaten the outside player more and extract more resources for any given bargaining protocol. However, when an alliance partner is larger, they demand a larger share of the alliance resources. Under $\mathbb{U}$, a player wants to ally with the strongest player possible. The larger total surplus generated outweighs the lower share received. $\mathbb{B}$ is the opposite. Players prefer to be with smaller partners because larger players are able to demand too high a share of the internal bargain. $\mathbb{B}$ operationalizes the assertion in Realism and its descendents that weaker states will allign against stronger states is the natural order of the state system. As Waltz aruges, “secondary states... flock to the weaker side; for it is the stonger side that threatens them."[54] Along these lines, Proposition 3.3 shows that $\mathbb{B}$ implies balancing behavior while $\mathbb{U}$ implies bandwagoning behavior.

**Proposition 3.3.** *Under scenario $\mathbb{U}$, the bandwagoning alliance ab is uniquely predicted in $\Gamma$. Under scenario $\mathbb{B}$, the balancing alliance bc is uniquely predicted in $\Gamma$.*

Another implication of Proposition 3.3 is that for $s_{ac}$, an alliance between the weakest and strongest players, is unlikely to be a stable state. For this to happen, it must be that utility is not changing in a uniform way with changes in the size of an alliance power. Namely, it must be that while if *a* substitutes an alliance with *c* for an alliance in *b*, the extra resources are not particularly empowering relative to the old alliance (war power concave in this part of the distribution), while if *c* substitutes an alliance with *a* for an alliance with *b* there is a large drop in the alliance's ability to make war threats (power convex in this part of the distribution). Alternatively, the additional gains for *a* from exploiting *c* within an alliance rather than the stronger *b* are greater than the loss of the alliance's ability to threaten war when *a* switches from *b* to *c*. On the other hand, the additional gains in war threats must outweigh the losses in exploitation for *c* when *c* switches from *b* to *a*. It is possible that a particular model specification will predict $s_{ac}$ in equilibrium. However, the lack of uniform comparative statics is suggestive that this particular alliance formation will be rare when only the static distribution of capabilities is considered.

Scenarios $\mathbb{U}$ and $\mathbb{B}$ give a partial characterization of when bandwagoning as opposed to balancing behavior occurs in the cooperative model. Since the cooperative models captures essentially any reasonable manner of transferring within a three-player set-up, within this context, the theoretical confidence of alliance predictions in these static cases is strong. One can find examples that seem to conform to both the $\mathbb{U}$ and $\mathbb{B}$ cases at the same time period in history. The 1892 Franco-Russian Alliance to balance the German Empire that

[54] See Waltz (1979), 127.

had become dominant on the continent would fit nicely into the $\mathbb{B}$ category. The weaker states (lower $r$) preferred to ally with each other, rather than joining with the strongest state, Germany, in demanding large concessions from the state left out of the alliance. The proportion of concessions going to Germany would have been too high for the potential ally to benefit. On the other hand, the understandings reached between the dominant European powers in dividing up Africa and China in the 19th century suggests a $\mathbb{U}$ structure. The powerful European states (high $r$) joined together to demand large concessions from the weaker (low $r$) Africans and Chinese. In this case there were less difficulties in finding accomodating sharing rules where all of the European states benefited.

However, since these characterization only apply to static models with players distinguished only by a single variable, they are more suggestive than definitive. One needs to begin to think about dynamic effects as well as the implications of many states with attributes more complex than a single resource variable before applying the theoretical predictions of this model to historical cases. However, the model does provide a framework to build on and to use to begin to disentangle these issues.

## 3.4 Dynamic Models

While the static models are useful for characterizing balancing versus bandwagoning behavior, they do not give an insight into the occurrence of conflict. In this section, I explore three different ways that dynamic change can be introduced into the static model. The first model begins as a two-player game at time 0, then a third player enters exogenously at time 1. The second model considers a game where in the first period, the three players are in a state $s$, but then a change occurs in the distribution of power that causes an alliance shift. This type of shift is unexpected at the time of alliance formation so players are stuck in their current alliance and sharing rule. This model is akin to directly adding the logic of Powell (2004, 2006) to the static three-player game introduced above. Finally, when resource shifts are anticipated, the current alliance state may alter in order to avoid war.

The dynamic games below all have the same fundamental setup. There is a one-time change following the first period, then the game is repeated infinitely. I only consider one-time changes as a simplification, but multiperiod changes would be an interesting extension. Players discount the future at a rate $\delta \in (0,1)$ per period. Players are assumed to not be able to commit to future bargains. War still results in the winner acquiring the resources of the loser(s) for all future periods. Hence, the value of war that eliminates all other players is scaled up by $\frac{1}{1-\delta}$. With these assumptions, all periods after the initial change are strategically the same as the static game.

This is not the case in the first period. It no longer makes sense to assume bargaining follows any particular protocol $\rho(\cdot)$ in the initial period. Instead, I will allow the players to make any bargain for that period that might avoid war. This ensures that when the model predicts war, it is not an artifact of an arbitrary bargaining protocol.

### 3.4.1 A third player as entrant

Consider a game that starts in period 0 with two players, $\alpha$ and $\beta$. Players are endowed with resources as before and assume that $r_\alpha > r_\beta$. That is, in comparison to the three-player model, $\alpha$ is larger than $\beta$, but a third entering player can be any size in relation to $\alpha$ and $\beta$. Players have war making ability as in the leading example so that if a player $i$ chooses war, $i$ wins with probability $q_{i0} = \frac{r_i}{r_i+r_j}$ and each player pays cost $\gamma_0 = \frac{r_\beta}{2(1-\delta)\left(\frac{r_\beta}{r_\alpha+r_\beta}\right)}$ which is a direct, present value modification for two players of cost in the three-player leading example. The winner captures the resources of the loser and is able to use them in war in future periods. To avoid war, the larger player, $\alpha$, proposes a utility transfer $\tau_t$ in each period $t$, which is positive if $\beta$ is to transfer to $\alpha$, and negative if $\alpha$ is to transfer to $\beta$.

The timing of the game in period 0 is

1. $\alpha$ proposes $\tau_0$ to $\beta$.

2. $\beta$ accepts the proposal and receives $r_\beta - \tau_0$ in period 0 or goes to war.

3. The game moves to period $t = 1$.

It is first useful to consider what this game would look like if there were to be no change in the system. Let $M_{i0}$ be $i$'s war value at 0, then

$$M_{i0} = \tfrac{q_{i0}\left(r_i + r_j\right)}{1-\delta} - \gamma_0$$

Since $\alpha$ has proposer power, $\alpha$ will propose a stream of payments to $\beta$ such that $\beta$ receives exactly $V_{\beta 0} = M_{\beta 0}$, while $\alpha$ receives $V_{\alpha 0} = M_{\alpha 0} + 2\gamma_0$. A sequence that satisfies this requirement gives $\alpha$ and $\beta$ this normalized amount in each period so that $\tau_t = r_\beta - (1-\delta) M_{\beta 0}$.

Now imagine that at the beginning of period 0 it is known that a third player, $\eta$, will exogenously enter the game at the beginning of period 1. There are two scenarios. First, period 0 concludes peacefully and alliances form as in the leading example. The value of this to a player $i$ beginning in $t = 1$ is $\frac{u_i(s)}{1-\delta}$. Alternatively, players $\alpha$ and $\beta$ go to war in period 0, and the winner is left with resources $r_\alpha + r_\beta$ to play the two player game with $\eta$.

Thus, player $i$ prefers war at time 0 for any transfer if

$$q_{i0}\left(r_\alpha + r_\beta + \delta V_{\alpha\beta 1}\right) - \gamma_0 > r_\alpha + r_\beta + \delta \frac{u_i(s)}{1-\delta} \tag{3.1}$$

$V_{\alpha\beta 1}$ represents the value function for a winning power with resources $r_\alpha + r_\beta$ bargaining with $\eta$. Thus if (3.1) is met for either player, war will result in period 0 followed be peaceful bargaining between the winning player and $\eta$.

**Example.** Consider the game above with the resource vector $r = \{50, 30, 20\}$, $\delta = 0.95$ and an initial two player game between $a$ and $c$. From Table 4, one can already calculate that the entering player $b$ will form an alliance with $c$, and the value that would have for $a$ beginning in period 1 is $u_a(s_{bc}) = \frac{40}{1-0.95} = 800$. Thus, for $a$, the RHS of (3.1) is $50+20+0.95(800) = 830$. On the other hand, if $a$ declares war in period 0, $a$ wins with probability $q_{a0} = \frac{5}{7}$ and pays cost $\gamma_0 = 280$. If $a$ wins, $a$ captures $r_a + r_c$ in period 0 plus the discounted value of future bargaining with $b$, $V_{ac1}$. $a$ would continue to be the larger power and therefore have proposer power, and the cost would be $\gamma_1 = \frac{30}{2(1-\delta)\left(\frac{70}{100}\right)} \approx 429$. Thus $a$'s war value would be

$$M_{a1} = \frac{\frac{7}{10}(70+30)}{1-0.95} - 429 = 971$$

Hence $V_{ab1} = 971 + 2(429) = 1829$. Therefore, the LHS of (3.1) is

$$q_{a0}\left(r_a + r_c + \delta V_{ac1}\right) - \gamma_0 = \frac{5}{7}\left(50 + 20 + 0.95(1829)\right) - 280 \approx 1011$$

Therefore, inequality (3.1) is satisfied in this case and $a$ prefers war in period 0 to any bargain in period 0.

This example demonstrates even if a two-player system may be otherwise peaceful, a player may go to war in order to prevent a power shift associated with a new player entering the system. The players go to war both to prevent an adverse alliance from forming, but also to gain resources in bargaining with the entering player.

An historical example of this type of war may be, as Tooze (2006) argues, Nazi Germany's decision to invade Russia upon the U.S.'s decision to aid Britain and enter the European system. In the summer and early fall of 1940 it became clear that Nazi Germany would not be able to knock out the British air force or bring England to its knees through bombing alone. Lacking air or naval superiority, the Germans could not bring its land forces to bear against the Britian. While Britain was not an offensive threat to Nazi Germany, a combined Anglo-American force using the Britsh isle as a launching ground was capable of defeating the Nazis. On September 5th 1940, the U.S. agreed to give England 50 destroyers in exchange for British bases in the Caribbean.[55] This signaled that the U.S. would eventually be entering the European system in the combination with the British. This signal would be borne out by the later Lend-Lease arrangement and the U.S.'s final entry into the war after Pearl Harbor.

Meanwhile, the European system had been reduced to two major players: Nazi Germany and Soviet Russia. They had reached an accomodation in trade and divying up Eastern Europe. However, faced with the prospect of a rapidly arming U.S. force entering the system, Germany's resources looked inadequate to fend off the joint Anglo-American forces. The combined 1938 GDP of Anglo-American controlled territory was 1,483 million PPP dollars in 1990 prices while the European territories controlled by Germany totalled only to 1,071 million dollars.[56] Furthermore, Russia, whom Germany was reliant upon for many raw materials, was likely to take advantage of any Anglo-American attack on Germany by switching to the allied side. Thus, by the logic of the model, Germany had a strong incentive to prevent this alliance switching and acquire resources by attacking Russia.

The logic behind Operation Barbarossa is by no means unique in the history of warfare. Another example may be the first Sino-Japanese War of 1894-1895, where Japan attacked China seeking resources to balance the firmer Russian entry into the the Far East that loomed with the beginning of construction of the Trans-Siberian railroad.[57] Even the American Civil War can be viewed through this lens with the South initiating war after the entrance of the Western territories into the American system primarily on the side of the

[55]See Churchill (1949), 414.
[56]See Tooze (2006), 384.
[57]Paine (2003) is suggestive of this type of argument.

free states.

### 3.4.2 Unanticipated one time shift in resources

Now I consider a version of the game that begins as a three-player system but experiences a shift in resources that changes the bargaining relations within the system. There are several ways to consider shifts in resources in a three-player game, but I will focus on one-time resource shifts after which resources remain constant in all future periods. I assume that knowledge of the shift occurs once players are already committed to an alliance and intra-alliance transfer scheme in period 0. This is the sense in which the shock to resources is unanticipated. However, the alliance and outside player are free to set the inter-alliance bargain. I label this bargain at 0, $\tau$. $\tau$ replaces the bargaining protocol $\rho$ in that period. Since the interest here is in identifying cases where war can occur, $\tau$ will be taken as the most favorable bargain achievable for a player considering war. Since resource shifts occur only once, $\rho$ still applies in future periods where there is no expected shift in resources. In period 1 onwards, the static game forms the stage game and there are no further shifts in resources. For simplicity, I drop the expectations notation and assume that players always get there expected value in each period.

Formally, I write the forced peaceful value functions[58] of a given player $i$ at $t = 0$ as

$$V_i^P(r_0, s(0), \rho) = v_i(r_0, s(0), \tau) + \delta V_i^P(r_1, s(1), \rho)$$

where $s(t)$ is the state at time $t$. This says that in peace, $i$ receives the value of being in state $s(0)$ at time 0 when the inter-alliance bargain it $\tau$. $s(0)$ is the solution to the static model when resources are $r_0$. Additionally, $i$ receives the discounted value of peace in the future when state is $s(1)$ and the bargaining protocol is $\rho$. The war value for the outside player is

[58]The value functions if players could not choose to go to war.

now the present value of the current period war value

$$W_i(r_0, s(0)) = \frac{w_i(r_0, s_{jz})}{1-\delta}.$$

An allied player $i$'s war value is the present value of the current period war value, which is also the bargain value when the outside player has all of the bargaining power

$$W_i(r_0, s(0)) = \frac{v_i(r_0, s(0), \rho = \gamma(s))}{1-\delta}.$$

Thus the full value function for a given player is

$$V_i(r_0 s(0), \rho) = \max\{v_i(r_0, s(0), \tau) + \delta V_i^P(r_1, s(1), \rho), W_i(r_0, s(0))\}.$$

War then occurs at $t = 0$ if for all transfers $\tau$, there exists a player $i$ such that

$$W_i(r_0, s(0)) > v_i(r_0, s(0), \tau) + \delta V_i^P(r_1, s(1), \rho).$$

The main qualitatively novel result in this model is that a state may initiate war even when relatively rising in power. In fact, as the next example shows, a player may initiate war even when the only resource change is that another player's resources are decreasing.

**Example.** Assume the functional forms of the leading example and let $r_0 = \{40, 35, 25\}$. In period 0, war values are

$$W_a(r_0, s_{bc}) \approx 633$$

$$W_b(r_0, s_{bc}) \approx 714$$

$$W_c(r_0, s_{bc}) \approx 486$$

The alliance-utility table in period 0 is

Table 7: Payoffs for the resource distribution: $r_a = 40$, $r_b = 35$, $r_c = 25$

| | $s_0$ | $s_{ab}$ | $s_{ac}$ | $s_{bc}$ |
|---|---|---|---|---|
| $u_a$ | 41 | 44 | 46 | 32 |
| $u_b$ | 35 | 38 | 27 | 41 |
| $u_c$ | 24 | 18 | 27 | 28 |

Now assume that a one-time decrease in $a$'s resources occurs so that $r_1 = \{34, 35, 25\}$. This generates the following alliance-utility table that will repeat infinitely from period 1 on:

Table 8: Payoffs for the resource distribution: $r_a = 34$, $r_b = 35$, $r_c = 25$

| | $s_0$ | $s_{ab}$ | $s_{ac}$ | $s_{bc}$ |
|---|---|---|---|---|
| $u_a$ | 34 | 37 | 39 | 26 |
| $u_b$ | 35 | 39 | 27 | 40 |
| $u_c$ | 24 | 18 | 27.8 | 27.6 |

In period 1, alliances shift from $s_{bc}$ in period 0 to $s_{ac}$. Thus the war inequality becomes

$$W_b(r_0, s_{bc}) \approx 714 > u_b(r_0, s_{bc}, \tau) + \delta V_b^P(r_1 s_{ac}, \rho) = 59 + 0.95(540) = 572$$

which holds, therefore $b$ prefers war in period 0.

This example is striking since this prediction is never made in the two-player model. A decrease in one player's resources can lead to war in the two-player model, but it is initiated by the declining player. Here, $b$ is rising relatively to $a$, but initiates war in order to avoid letting $c$ switch alliances. This is an important result because there are many historical cases where the rising power seems to have been the aggressor. For instance, Germany seems to have been the rising power at the beginning of World War II, but was clearly the aggressor.

Beyond this qualitative result, one can use this model to make some quantitative comparisons between this model and the two-player model of war. As a comparison, I define a scenario labeled $F$ for "fixed" where the initial alliance structure, $s(0) \neq s_0$, is fixed from period 0 on. The $F$ scenario then corresponds to the equivalent two-player case. This is because the allied players are effectively "stuck" together and act as a single player for war making decisions. I will call the scenario where players are free to switch alliances, $D$, for dynamic. The only difference between $D$ and $F$ is that there is no alliance switching in $F$. Therefore, the war condition in $F$ at $t = 0$ is if for all transfers $\tau$, there exists a player $i$ such that

$$W_i(r_0, s(0)) > u_i(r_0, s(0), \tau) + \delta V_i^P(r_1 s(0), \rho).$$

The following example illustrates how a shifts in resources impact war decisions in the leading example.

**Example.** (i): Consider the setup of a dynamic version of the leading example with $r_0 = \{50, 30, 20\}$. From earlier analysis, this leads to state $s_{bc}$. Let $r_1 = \{50 + g, 30 - g, 20\}$ where $g \in (0, 30]$. Since this does not change the alliance structure of the system, $s(1) = s_{bc}$. $b$'s war value at time 0 is $W_b(r_0, s_{bc}) = \frac{u_b(r_0, s_{bc}, \rho = \gamma(s_{bc}))}{1-\delta} \approx \frac{31}{1-0.95} = 620$. The most $a$ can transfer to alliance $bc$ at $t = 0$ is that periods level of resources for $a$ which is 50. After invoking the sharing rule in the leading example, this implies that $u_b(r_0, s_{bc}, 50) = 61$. Thus $b$ will initiate war if

$$620 > 61 + 0.95 V_b^P(r_1, s_{bc}, \rho)$$

$$V_b^P(r_1, s_{bc}, \rho) < 588$$

Given this, one can then search for the minimum value of $g$ that satisfies this inequality. Here, the inequality is satisfied if $g \gtrsim 7$.

In (i), there is no alliance shifting. This means that the same conclusion would be found in the $F$ case and therefore, the three player game could be treated as a two-player game without loss. (ii) demonstrates a case where alliance shifting does occur leading to different conclusions under $F$ and $D$.

(ii) Now let $r_0 = \{40, 35, 25\}$ and $r_1 = \{35, 45, 20\}$. War values are already calculated for time 0 in the previous example and the alliance-utility table for period 0 is the same as Table 7. Table 9 gives the alliance-utility table for periods 1 on.

Table 9: Payoffs for the resource distribution: $r_a = 35$, $r_b = 45$, $r_c = 20$

| | $s_0$ | $s_{ab}$ | $s_{ac}$ | $s_{bc}$ |
|---|---|---|---|---|
| $u_a$ | 35 | 37 | 42 | 27 |
| $u_b$ | 46 | 49 | 36 | 52 |
| $u_c$ | 18 | 14 | 22 | 21 |

Under $F$, the state must remain $s_{bc}$ in period 1, but under $D$, alliances shift in period 1 so that $s_{ac}$ forms in period 1. A comparison of these two scenarios illustrates how alliance shifting affects conflict above and beyond the two-player model. While $a$ does worse in every state after $b$'s relative rise, $a$'s losses are mitigated by switching from isolation to forming a balancing alliance with $c$, to the point that $a$ is better off having given up it's hegemonic, but isolated position. Under $F$, $a$ is worse off after $b$'s rise. $c$'s is indifferent over the name of its alliance partner, but cares about its relative strength in demanding concessions. Therefore, under $D$, $c$ maintains its positions, while under $F$, the rise of $b$ hurts $c$. Interestingly, only $b$ is hurt by its own rise as $c$ switches into an opposing alliance.

The implication is that under $F$, one must examine both $a$ and $c$'s incentive for war, but under $D$ it is actually $b$ who might initiate war in period 0 while still in a committed alliance with $c$.

Under $F$, $a$ prefers war if

$$W_a(r_0, s_{bc}) \approx 633 > u_a(r_0, s_{bc}, \tau) + \delta V_a^P(r_1 s_{bc}, \rho) = 100 + 0.95(540) = 613$$

so that $F$ initiates war for the same preventive reasons as in the two-player game.

Under $D$, war occurs if

$$W_b(r_0, s_{bc}) \approx 714 > u_b(r_0, s_{bc}, \tau) + \delta V_b^P(r_1 s_{ac}, \rho) = 59 + 0.95(720) = 743$$

which is not satisfied, therefore war does not occur under $D$. While in some cases the impact of alliance shifting can be so severe as to cause war, in this case the three-player game predicts peace, whereas a two-player version, where an alliance between $b$ and $c$ was treated as a single player, war is predicted.

This example demonstrates the effect of alliance shifting on war in a particular case. Proposition 3.4 gives a general characterization of the effect of alliance shifting on conflict when a single player increases in resources and the structure of the game falls into either $\mathbb{U}$ or $\mathbb{B}$.

**Proposition 3.4.** *The following statements partially characterize the effect of alliances on the occurrence of war for a one-time, one-player positive resource shock in D in comparison to F:*

*Under* $\mathbb{U}$ *so that* $s_{ab}$ *is uniquely stable:*

*U1: If c experiences the resource shock and* $r_{c1} < r_{b1}$*, then a and b initiate war under the same conditions as in the F case.*

*U2: If c experiences the resource shock and* $r_{c1} > r_{b1}$*, then b initiates war under more* *initial conditions and shock sizes then in the F case. The effect on a's instigation of war is ambiguous.*

*U3: If a or b experiences the resource shock, then c initiates war under the same conditions as in the F case.*

*Under $\mathbb{B}$ so that $s_{bc}$ is uniquely stable:*

*B1: If a experiences the resource shock, then b and c initiate war under the same conditions as in the F case.*

*B2: If b or c experiences the resource shock and $r_{b1} < r_{a1}$ or $r_{c1} < r_{a1}$ respectively, then a initiates war under the same conditions as in the F case.*

*B3: If b or c experiences the resource shock and $r_{b1} > r_{a1}$ or $r_{c1} > r_{a1}$ respectively, then a initiates war under less initial conditions and shock sizes then in the F case. The originally allied player instigates war under more initial conditions and shock sizes then in the F case. The rising power may initiate war in D.*

Assuming that $u_a(r_1, s_{ac}, \rho) \geq u_a(r_0, s_{ab}, \rho)$ eliminates the ambiguity in U2 in favor of implying that the availability of alliance switching makes war more likely in this case. However, it is difficult to see what type of simple assumption could eliminate the comparative static ambiguity from the B3 case.

The B3 case lends credence to the complexity and difficulty associated with hegemonic transition. Balancing behavior makes the previous hegemon less likely to initiate war than predicted in the two-player model, however, the rearrangement of alliances can cause war as the previous alliance breaks up.

### 3.4.3 Anticipated one-time shift in resources

In the previous section, the change in resources happens after the alliances have already been determined. However, if the change was anticipated before the alliances formed in period 0, then the anticipation of dynamic change may impact the formation of alliances themselves. For instance, as shown above, in the case where $r_0 = \{40,35,25\}$ and there is an anticipated shock such that $r_1 = \{34,35,25\}$, if $c$ were to ally with $b$ in period 0, war would be assured, and $c$ would receive its war value of $W_c(r_0, s_{bc}) \approx 486$, yet if $c$ allied with $a$ period 0, $c$ would receive $u_c(r_0, s_{ac}, \rho) + \delta V_c^P(r_1 s_{ac}, \rho) = 27 + 0.95(556) \approx 555$, thus $c$ would prefer to ally with $a$ in period 0 and avoid war. Hence, one needs to know not just the distribution of military capabilities in a system, but also the anticipated direction of dynamic change in order to predict alliance formation accurately. Indeed, what may seem like bandwagoning from a static standpoint may be necessary to avoid war from a dynamic perspective.

As in the above example, the dynamics of the distribution of power are important when the anticipated shift will lead to war. If no war would result in the unanticipated case, then it is enough to know the current distribution of power in order to predict which alliances will form in a given period. It is the pressure of an inefficient war that causes the change in the current period's alliance outcome. For instance, one might imagine that Australia may be able to receive short term bargaining gains today by balancing against the U.S. with a rising China. However, Australia would want to switch and balance against China in the future if it overtook the U.S. in power. This could lead to China initiating a general war while Australia is still committed to an alliance. In order to prevent this, Australia would prefer not to enter an alliance with China in the first place.[59]

[59] In a model where shifts in resources are persistent as in Chapter 1, the initial shock to the trend may be unanticipated, but future shifts in power are anticipated. Since China's rise relative to the U.S. is a persistent trend, exploring the impact of persistent shifts in a multilateral model would be a particularly interesting extension.

Taking this a step further, the effect of shifts in the distribution of power on alliance formation may be particularly important in trying to understand the general lack of hard balancing against the U.S. after the end of the Cold War. While it is possible to imagine a coalition including China that may effectively balance the U.S. and perhaps even provide short run bargaining gains for those currently chafing under U.S. hegemony, China's rise and the potential for conflict with the U.S. would disincline states against this type of balancing. It is the relevant decline of the U.S., despite its dominant position in the distribution of power, that discourages military balancing. Recent attempts to tighten the economic relationship between the U.S. and the European Union may appear to have the elements of a bandwagoning alliance in the static sense, but could actually be dynamic balancing against a rising China.

## 3.5 Conclusion

The static three-player bargaining model of war allows one to apply the insights of the bargaining model of war to alliance formation as well investigating the implication of shifting alliances on the occurrence of war. If the increase in inter-alliance bargaining power associated with allying with a stronger power dominates the decrease in intra-alliance share associated with having a stronger ally, the strongest states will bandwagon against the weakest. If the opposite is true, the weakest states will balance the strongest. An alliance between the strongest and weakest state is only possible in the static model if neither assumption holds at all levels of resources.

Allowing for a dynamic element introduces two novel mechanisms for war. The anticipated exogenous entry of a third state into a two-player system can cause the initial two states to go to war in order to prevent an adverse alliance with the new state and to capture resources to use in bargaining with the entrant. When the alliance structure is fixed and resource shock occurs, a state may initiate war even when that state is rising in relative power. Furthermore, the effects of the anticipated alliance shift can both exacerbate or ameliorate the potential for war relative to the two player model. If the alliance structure is not fixed when a future shift in resources occurs, alliance formation need not correspond to the static prediction.

The model in this chapter can be extended in three directions. First, it is subject to all of the extensions to the two-player model found in the literature. Adding incomplete information, strategic militarization, first strike advantages, and multiperiod changes in resources may have modified effects in the three-player model. Second, while much of the intuition behind alliance formation can be handled parsimoniously in a three-player model, adding more players would improve the empirical applicability of the model. However, increasing the number of fully strategic states quickly leads to tractability issues. Another approach might be to add a number of non-strategic "fringe" states or minor powers. Competition by the strategic major powers over the resources held by the non-strategic minor powers could add rather interesting dynamics to the system and address historical cases like the U.S. involvement in Vietnam. Finally, the complexity of the underlying structure of the game could be increased to better mirror reality. Examples may include different types of economic resources with varying degrees of complementarity, differing military resources like land and sea power, disparate compatibility in terms of the political and social life between states, and including an explicit network structure of the international system.

# A Alternative Bargaining Protocols for Chapter 1

In general, preventive war occurs when changes in bargaining space of future periods imply a backward induction solution of $w$ in the current period. In the Chapter 1 model under bargaining protocol $\mathbb{S}$, there exists only one type of change to the future bargaining space. Namely, the set of bargains that are minmax compatible at time $t$ is shifting as $t$ increases so that the upper bound on $x$ is falling. This means that $S$ is doing worse and worse in future bargains, so $S$ may choose to lock in its current war value if this deterioration is significant enough. However, this is not the only change to the bargaining space that can cause preventive war. I explicate some of these other changes using some alternative bargaining protocols.

By using the protocol $\mathbb{S}$, the analysis in the previous section picked out the specific point in the minmax compatible bargaining space at which $S$ did the best. This is because $S$ had all of the bargaining power to determine the allocation of the pie, $B$, that was in excess of the minmax constraints. The only way in which $R$ gains in bargaining power is by the increase in its minmax value over the length of the shock. Consider the following two alternative bargaining protocols: $\mathbb{R}$ and $\mathbb{U}$. Under $\mathbb{R}$, $R$ makes a take-it-or-leave-it offer in each period. Under $\mathbb{U}$, for all periods $t < T$, $S$ makes a take it or leave it offer and in all periods $t > T$, state $i$ receives $(1-\delta)M_{iT} + \frac{d}{2}$. $\mathbb{R}$ represents the case where the assignment of bargaining power is the most averse to the prospect of peaceful bargains. $\mathbb{U}$ represents the case where bargaining power is even after $R$'s rise is complete.[60] This seems to be the natural case where along with $R$'s inability to commit to not exercising future bargaining strength through its minmax value, $R$ also cannot commit to giving $S$ a privileged bargain outcome when $R$ is no longer worried that $S$ might go to war.

[60] I allow $S$ to make the offers when $R$ is rising only for ease of exposition. The value of the bargain after $T$ is a robust outcome when bargainers do not face changing outside options. It is both the one period Shapley value and the Rubinstein outcome when bargaining over the pie when the proposer is random and there is a discount value $\beta \to 1$ when waiting for offers within a period $t$.

First consider the bargaining protocol $\mathbb{R}$ for a given $p_0$, $\theta$, and shock length $T$. Since $R$ has the power to make take it or leave it offers in all periods $t \geq T$. I assume that $R$ must play a SPE, therefore $R$ offers $x_t = (1-\delta)M_{ST}$ in all periods $t \geq T$, This means that starting in period $T$, $S$ values peace at exactly its war value or minmax value $M_{ST}$. That means that in order for a peaceful bargain to exist, it must be that $S$ prefers the discounted series of bargains in period $t < T$ plus its discounted war value at $T$ to its war value at any time $t < T$. Since the most $R$ can give $S$ at $t < T$ is $x_t = 1$, the following condition is necessary and sufficient for war.

**Proposition A.1.** *War occurs in any SPE of the game, $\Gamma$, under bargaining protocol $\mathbb{R}$ at $t = 0$ if and only if inequality (1.4) holds.*

This is quite a striking result since now (1.4) implies war for any shock length $T$. Essentially, this follows since $\mathbb{R}$ implies that the best $S$ can do in a future bargain is equivalent to its war value in the future. Proposition A.1 demonstrates that a lack of commitment in proposer power can cause war in situations where shifts in the distribution of power alone are not enough to generate war.

Neither $\mathbb{S}$ nor $\mathbb{R}$ is a realistic way of modeling a bargaining process, but they are useful for purposes of exposition. $\mathbb{U}$ is not necessarily realistic either, but it better approximates what one should expect in periods $t > T$. Here the bargain outcomes in periods after $T$ simulate the case where bargaining power is even when there are no state changes. The effective bargaining spaces in periods after $T$ are reduced to a single point in the middle of that period's minmax compatible bargaining space. $\mathbb{S}$ and $\mathbb{R}$ also reduce the effective bargaining spaces to single points, but at the extremes of the minmax compatible bargaining space. Proposition A.2 gives a necessary and sufficient condition for war under $\mathbb{U}$.

**Proposition A.2.** *For $\theta$ large enough to satisfy* (1.4)*, war occurs in any SPE of the game, $\Gamma$, under bargaining protocol $\mathbb{U}$ at $t = 0$ if and only if*

$$\theta > \frac{1}{\delta}\left(1+\frac{d}{p_0(1-d)}-\frac{d\delta^T}{2p_0(1-d)}\right)^{\frac{1}{T}} \tag{A.1}$$

Clearly (A.1) is a weaker constraint on the occurrence of war than (1.5). The term $-\frac{d\delta^T}{2p_0(1-d)}$ dictates the range of $\theta$ that leads to war under bargaining protocol $\mathbb{U}$, but not under $\mathbb{S}$. The effect of switching from bargaining protocol $\mathbb{S}$ to $\mathbb{U}$ is particularly dramatic when $T = 1$. Since this corresponds to the LRM case, this type of bargaining protocol can reduce the high barrier imposed by the LRM on empirical cases.

Table 10 compares the minimum $\theta$ that causes war under all three bargaining protocols outlined here using the parametrization of the leading example.

Table 10: PSM 2: $p_0 = \frac{1}{3}$, $\delta = 0.96$, $d = 0.2$

| $T$ | 1 | 2 | 3 | 4 | 6 | 8 | 10 | 12 | 14 |
|---|---|---|---|---|---|---|---|---|---|
| Min $\theta$ under $\mathbb{S}$ | 1.82 | 1.38 | 1.26 | 1.20 | 1.14 | 1.12 | 1.10 | 1.09 | 1.08 |
| Min $\theta$ under $\mathbb{U}$ | 1.45 | 1.23 | 1.17 | 1.14 | 1.11 | 1.09 | 1.08 | 1.08 | 1.07 |
| Min $\theta$ under $\mathbb{R}$ | 1.07 | 1.07 | 1.07 | 1.07 | 1.07 | 1.07 | 1.07 | 1.07 | 1.07 |

Shifts in the distribution of power shift minmax values in the same direction. Different minmax compatible bargaining protocols assign values within these minmax values in various ways. This section has shown that combining the two can lead to sharper conditions on war. One can also imagine minmax values shifting individually, contracting or spreading apart. Furthermore, if the per period bargain size is change, the entire bargaining space can stretch or contract. It is possible to implement all of these types of shifts as extension to the model in Chapter 1. Combining them with particular bargaining protocols such as $\mathbb{S}$, $\mathbb{R}$, or $\mathbb{U}$ can lead to war without shifts in the probability of victory.

# B A General Two-Player Model of Preventive War

## B.1 Model

### B.1.1 State Variables and Actions

Consider a bargaining model of war set in a stochastic game labeled $\Gamma$. Time is discrete and infinite starting at 0 so that $t \in \{0, 1, 2, ...\}$. There are two players generically denoted $i$ and $j$. Both players discount the future at a common rate $\delta \in [0, 1)$.

Call the state $k \in K$ where $K$ is the set of all possible states. I assume that $K$ has finite elements. A sequence of states beginning at time 0 and ending at $t$ will be referred to as $k^t$. I will refer to the state at a particular time $t$ as $k_t$ or $k_t^\tau$ to emphasize that it comes from a particular sequence ending in period $\tau \geq t$. When referring to a state variable that is particular to player $i$, I will call it $k^i$ for the entire sequence and $k_t^i$ at a time $t$. The state variable can be a very rich concept and is meant to represent the entire relevant state of the world at the time of the war decision. Examples of relevant states might be the military resources of a player, the beliefs of a player about another player's military strength, the random outcome of an uncertain event, or the results of a game played within the time period $t$ that determines the state relevant to the war decision at the end of period $t$.

If war has not occurred yet, each player has two possible actions, $\{w, d\}$, where $w$ means starting a war and $d$ means declining to start a war. In order to avoid unnecessary complication of simultaneous choice, I make the assumption that the players choose their action in an arbitrary order. Once one player chooses war the game ends. There are no advantages to striking first within a period.[61]

I model war as a one shot process that ends the game. I do this since the model is

[61] This does not mean this model cannot be used to model first strike advantages. One method would be to give player $i$ a higher value of war in odd periods and player $j$ a higher value in even periods. If $i$ does not attack in an odd period $i$ gives up its first strike advantage and similarly for $j$ in even periods.

primarily interested in determining when a war will occur and not the nature of the war and peace process itself. The valuation players place on choosing war includes all future foreseen implications of that decision. The fact that war ends the game means that whether or not war has ever occurred is always one of the elements of the state variable. Since no action occurs after war, I assume throughout unless explicitly stated that we are in a state where war has not yet occurred.

### B.1.2 The Bargaining Space

The generic critique of the bargaining model of war is that any model that predicts war in a given situation must explain why the two players did not have a mutually beneficial bargain available to them in order to avoid that war. I represent this critique as a variable $b \in B$ where $b$ represents the sequence of bargain outcomes for both players in all possible future states and $B$ is the set of all possible sequences of this type. The bargain outcome at time $t$ is $b_t$. I will sometimes find it convenient to refer to $i$'s part of the bargain as $b^i$ and in period $t$ as $b^i_t$.

In order to fix the idea of what $b$ and $B$, consider the following example. for all analyzing a situation where two countries are dividing a piece of land of size 1 in each period. Assume there are no other assets to bargain over. Let $b^i_t$ represent how much of the land $i$ holds at $t$. Imagine there are two states $k^{'}$ and $k^{''}$ that may occur in each period with probability 0.5. A particular $b$ might specify that in all periods $0 \leq t \leq 50$, if the state is $k^{'}$, then the bargain is $(b^i_t, b^j_t) = (0.75, 0.25)$ and if the state is $k^{''}$, the bargain is $(b^i_t, b^j_t) = (0.25, 0.75)$. Then for all $t > 50$, $b$ might specify regardless of the realized state, the bargain is such that $(b^i_t, b^j_t) = (0.5, 0.5)$. $B$ is the set of all such possible sequences, so in this case $B$ would be all bargains such that $0 < b^i_t + b^j_t \leq 1$ for all possible sequences of $k_t$.

Any possible bargain process must, in the end, specify the outcomes in all states for a given state sequence, possibly as a result of actions taken. The set $B$ captures all of these possible outcomes. How $b$ is chosen for a particular case is left abstract but can be thought of as the outcome of a games within each period $t$ or the actions of a social planner. I leave $b$ abstract for three reasons. First, no specific bargaining process can capture the reality of bargaining in the real world. Any particular process is open to the critique that it is too restrictive and war results in the model because of the arbitrary process chosen. By showing when war will occur considering the entire bargaining set $B$, the model makes a much more compelling case. Second, it is actually easier to show the sufficient condition in this abstract setting than for any but the simplest bargaining processes. Third, by considering this general set of bargains, I can take any bargaining process considered by a particular model that fits within our framework as a sub-case.

Since many models consider only a subset of $B$ as possible, I also consider allowing for general restrictions on the bargaining space in order to capture these models as sub-cases. Let there be the variable $\phi \in \Phi$ where $\phi$ is a subset of $B$ and $\Phi$ is all possible subsets.[62] Let the constrained set of possible bargains at period $t$ be $\phi_t$, so that in a given period, the relevant bargain is some value $b_t \in \phi_t$. If one player chooses $w$, neither player receives the bargain value. If both choose $d$, then the bargain $b_t$ occurs and the players transition to the next period of the game.

The bargains one wants to consider for a particular specification of our game may depend on the state in that period. When this is the case, I will make this explicit by writing $\phi_t(k_t)$ or $b_t(k_t)$. Of course, the bargain itself can be thought of as a state for purposes of the war decision. However, it is special here in that we want to consider all possible bargains that might occur when searching for a sufficient condition for war.

[62]The most extreme cases of this are where bargaining is not allowed. In these cases $\phi$ is a singleton value corresponding to the status quo. Another example of a restriction would be when two countries are bargaining over external land, but the actual land the countries possess is not on the bargaining table. This might be a reasonable restriction corresponding to the case of a domestic political constraint where losing a country's land will result in severe political consequences for a decision maker. $B$ can be thought of as the "natural" bargaining space for a problem. $\phi$ is the restricted set that arises from the modeler's assumptions.

### B.1.3 Timing

Each time period $t$ has the following subdivision. First, players observe the state $k_t$. Next players observe the bargain value $b_t$. The values of $k_t$ and $b_t$ are potentially determined by a more complex process within the game, but the important point is that their value is known prior to the war action phase. After observing the state and bargain values, players choose action ($w$ or $d$) in a random order. If either choose $w$, the game ends in war. If both players choose $d$, players receive bargain values according to $b_t$ and move to period $t+1$.[63]

### B.1.4 The Transition

The game $\Gamma$ evolves from period to period according to the Markovian transition function $q(k_{t+1}|k_t, b_t)$ when both players choose $d$. The game ends in period $t$ when at least one player chooses $w$. Using this notation, $q$ is the probability density that the next period's state is equal to $k_{t+1}$. The transition is dependent on today's bargain since today's bargain may endogenously affect tomorrow's state in some specifications (Leventoglu and Slantchev (2007) is such a model). $q$ may be deterministic for some specifications.

It is important to note that the state $k_t$ is summary of everything that is relevant for the war decision at time $t$. One may write a model in this framework where players enter time $t$ with some preliminary state, $s_t$, then take optimal non-war decision actions from some action set $A_t$. The combination of the preliminary state and realized actions would then make up $k_t$. $k_t$ in turn determines the bargain that is reached at time $t$ according to the particular bargaining process $b$ that we are considering. $k_t$ and $b_t$ then go on to determine the probability that state $k_{t+1}$ occurs. The probability of $k_{t+1}$ may result from a combination

[63] The time period ends simultaneous to receiving the bargain value since this could potentially have an endogenous affect on the state variable and therefore should be properly modeled as separate periods. The modeler chooses the length of time represented by the discrete period $t$. If there is a foreseeable substantial change to the state of the world, it is an arbitrary restriction to not allow players the possibility of war both before and after the change.

of the probability of a preliminary state $s_{t+1}$ and then the probability of actions $a_{t+1} \in A_{t+1}$ that ultimately determine the relevant state $k_{t+1}$. However, since the optimal action or mixing of actions is known once $s_{t+1}$ is determined, one can summarize this probability with the single equation $q$, dependent only on $k_t$ and $b_t$.[64] [65]

For an entire sequence $k^t$, the transition function $q$ implies a probability of sequence $k^t$ dependent on the initial state $k_0$ and particular bargain $b$. Call this probability $\rho(k^t|k_0,b)$.[66]

### B.1.5 The Value of War and Peace

If either player chooses $w$ in a given period it implies that both players must go to war. One can then think of a player's war value as the player's minmax value. Therefore, I will say that if war occurs in state $k_t$ then player $i$ will receive a payoff of $M^i(k_t) \in [-G,G]$ where $G$ is a (possibly very large) positive number so that $M^i(k_t)$ is bounded. As with $q$, unless otherwise noted we will take $M^i(k_t)$ as known.

If neither player chooses $w$ when the state is $k_t$ and the bargain is $b_t$ then the continuation

[64] One might worry that we have eliminated any bargaining process that is not the result of a Markov strategy. However, note two points. First, I allow actions to constitute states in this model. Second, I allow time to be an aspect of the state. These two points mean that Markov refinements are largely irrelevant for our model.

[65] In some models, the underlying game that is considered may have multiple equilibria for a given $k$ even without bargaining or within a given bargain. In this type of model $k$ and thus $q$ are indeterminate. From the perspective of Appendix B, these equilibria represent completely different specifications. However, one could check all of these specifications systematically. The model only considers one of these equilibria possibilities for $k$, then considers all possible bargains, $b$, for that single equilibria. This then solves whether that particular equilibrium leads to war or not. To solve for all possible equilibria for $k$, one would then have to apply the technique here to all other equilibria, just as for each particular equilibrium for $k$ I solve for all possible bargains, $b$. In other words, in the algorithm of this section, solving for all possible $b$ for a given $k$ is the inner loop, then the outer loop is solving for all possible $k$. In certain cases, it may be obvious that only one or a few equilibria need to be checked in order to determine if the sufficient condition for war is met for all specifications.

Also, it may be that in many cases where this type of multiple equilibria arise that something is missing from the model such as a reasonable coordinating device. For instance, if the players are leaders of a country, it may be that there are multiple growth path equilibria, say one corresponding to normal growth and another to a poverty trap. In the context of analyzing war, why does this model not specify more clearly the mechanism driving coordination to one of these equilibria or at least specify a probability distribution over the equilibria?

[66] Let $\rho(k^0|k_0,b) = 1$.

value of the game is $V^i(k_t, b_t)$ for player $i$. Sometimes I write $V^i(k_t, b)$ when needing to emphasize that I am writing the value function for a particular bargain sequence. One can think of player $i$'s current period utility in the absence of war as $u^i(k_t, b_t)$ where $u^i$ is nondecreasing in $b_t^i$ and where it is assumed that $u^i \in [0, G]$ so that it is bounded and hence $V^i(k_t, b_t)$ is bounded as well.[67]

For a given bargain $b$ one can write $V^i(k_t, b_t)$ recursively. However, while I assume for $V^i(k_t, b_t)$ that war does not occur today, it is not natural to make that assumption for the next period. To do so would give an inaccurate value for not going to war today since it very well may occur in the future. Let $K_t^W(b)$ be the set of states that result in war under bargain $b$ at time $t$ under rationality and let $K_t^P(b)$ be the set of states that result in peace under rationality. Importantly, whether war occurs at $t$ or not depends on the entire sequence $b$ starting in period $t$ since the future of bargains are important to today's valuation of war and peace. By under rationality I mean what a rational agent $i$ would choose given that $i$ knows $M^i(k_t)$ and $V^i(k_t, b_t)$ exactly. Throughout I assume that agents are rational. That is, agents have complete and transitive preferences. However, I do not assume that as modelers one knows the value of $V^i(k_t, b_t)$, which can be quite complex.

One can then write the recursive formulation of $V^i(k_t, b_t)$ as

$$V^i(k_t, b_t) = u^i(k_t, b_t) +$$
$$+\delta[\textstyle\sum_{k_{t+1} \in K_{t+1}^W(b)} q(k_{t+1}|k_t, b_t) M^i(k_{t+1}) + \sum_{k_{t+1} \in K_{t+1}^P(b)} q(k_{t+1}|k_t, b_t) V^i(k_{t+1}, b_{t+1})]$$

The first term on the RHS is the value of the bargain today. The terms within the square bracket are the value of the game beginning tomorrow which is discounted by $\delta$. The first term in the square brackets is the expected value of the states that lead to war next period.

[67] It is convenient to make the lower bound of $u^i$ zero. Any utility function that is bounded above and below over the state space (which is assumed finite) can be mapped into $u^i$.

The second term in the square brackets is the expected value of the states that do not lead to war next period.

My general assumption is that the player $i$ knows the value of $V^i(k_t, b_t)$, but we as modelers do not. For complex specification of the states even calculating the value of $V^i(k_t, b)$ for a single bargain $b$, could be quite difficult. Trying to do this for all $b \in \phi$ would be extremely difficult.[68] Thus, in the characterization section we focus on finding an upper bound of $V^i(k_t, b)$ such that $b \in \phi$. Furthermore, since I am using bounds, our analysis robust to misspecification in the bargaining game that the players are engaged in.

### B.1.6 Equilibrium

For a given bargain $b$, whether player $i$ takes action $w$ or $d$ is a simple optimization exercise. Choose $w$ if for $b$, $M^i(k_t) > V^i(k_t, b_t)$ and choose $d$ otherwise.[69] I am, however, interested in when war will occur for any allowable bargain and not just a particular bargain. Thus I will say that war occurs in period $t$ if it is reached and

**Condition 1.** $\exists\, i$ such that $M^i(k_t) > V^i(k_t, b_t)\ \forall\, b \in \phi$.

Otherwise, peace occurs in period $t$ and some bargain $b_t \in \phi_t$ is agreed upon in that period.

Many models specify $b_t$ as the result of an action taken in a period $t$. Consider such a model $\gamma$ which can be modeled as a sub-case of $\Gamma$. In $\gamma$ the actions taken that determine the bargain $b_t$ must produce a $b_t$ that is within some space $\phi_t$. Bargains today may affect bargains in the future so actions today may limit the entire sequence of future bargains to some set of sequences $\phi$. Since I assume that players optimize in each sub-game beginning

[68]One might worry that players themselves would not be able to calculate accurately their own valuation of such a complex object. This is not a problem as any bounded rationality would be included in the valuation of $V^i(k_t, b_t)$ itself. I am seeking to model the player's actual decisions, not the hypothetical best decision they could make.

[69]I assume players choose peace at equality.

at all periods $t$, i.e. players choose war for a given $b$ if $M^i(k_t) > V^i(k_t, b_t)$ and $d$ otherwise, when applying results from our general model I am inherently only considering Sub-game Perfect Equilibrium (SPE) of a particular game like $\gamma$. This is, however, a natural equilibrium refinement of Nash Equilibrium when players lack a commitment technology, which is the assumed setting in the international system. Condition 1 implies that all SPE lead to war in the case of $\gamma$. That is, there is no credible sequence of bargains that avoids war.

One could place reasonable restrictions on the types of bargain strategies players $i$ and $j$ can employ in $\gamma$. For instance, one might say that players are limited in their bargain strategies to those that are renegotiation proof in some sense. This would serve in our model as an exogenous restriction on the bargain space represented by $\phi \subset B$. Such a restriction would simply eliminate potentially peaceful equilibrium from consideration, thereby increasing the bite of our sufficient condition for war.

## B.2 Characterization

Finding when Condition 1 is satisfied is the central challenge of characterizing the bargaining model of war. The problem in doing this is that $M^i(k_t)$, $V^i(k_t, b_t)$, and $\phi$ can be quite complex objects for a given specification. Specifically the value function $V^i(k_t, b_t)$ may be far from obvious since it depends on an infinite number of bargain and war decisions. The typical strategy in the literature is to take $M^i(k_t)$ as exogenous according to some process specified in the model. This is reasonable since we are interested in the actual war decision as opposed to an analysis of the war process itself. Furthermore, this exogenous process can be seen as arising from optimally taken actions and thus can be used in models with endogenous changes so long as the outcomes can be specified as equilibrium outcomes.

I will also take $\phi$ as given. I will emphasize this exogenous nature by writing $\phi^{ex}$. This is simply to say that a given bargaining model of war might in fact take certain limitations on bargaining as given. Bargaining models frequently do this by specifying a particular bargaining process. I simply make this assumption explicit and at the same time more flexible since any or no bargaining restriction is possible using this specification.

Using the exogeneity of $M^i(k_t)$ and $\phi^{ex}$ one can proceed to find a sufficient conditions for war. The idea is that it may be difficult to immediately discern the value $V^i(k_t, b_t)$ exactly, but one may be able to bound it. The tighter this bound, the better the sufficient condition.

Let $\Xi(k_t)$ be an upper bound on the mutual value starting in state $k_t$. I will follow Powell (2004) and assume that the game is specified well enough that the maximum mutual value is known. Let $b_t^{\#} = \arg\max_{b_t \in \phi_t^{ex}} [V^i(k_t, b_t) + V^j(k_t, b_t)]$ so that $b_t^{\#}$ is a bargain that maximizes the mutual value of not going to war today and thus our upper bound is $\Xi(k_t) = V^i(k_t, b_t^{\#}) + V^j(k_t, b_t^{\#})$, the maximum possible expected mutual value for the two players.

I now proceed to build toward the main result, Proposition B.1, which proves a sufficient condition using a rather tight bound on $V^i(k_t, b_t)$. First, consider only looking at the current period's minmax values. I arrive at the following sufficient condition for war which we label $SC(0)$.

**Lemma B.1.** *($SC(0)$ sufficient condition for war) If state $k_t$ is reached and $M^i(k_t) > \Xi(k_t) - M^j(k_t)$, then war occurs at $t$.*

The $SC(0)$ condition holds only if the mutual value of war exceeds the value of any peaceful bargain. Since one typically assumes that war is inefficient compared to some potential bargain, one can think of these cases arising primarily in the context of a limited bargaining space or private information.

Having shown $SC(0)$, I now prove a broader sufficient condition, $SC(1)$. This turns out to correspond to a modification of the inefficiency condition found in Powell (2004). This condition is less general than Powell's since it only applies to bargaining models of war as specified here, but is applicable to models with private information in that I allow the condition in Lemma B.1 to lead to war. First, I must introduce the following notation:

*Notation* 1. $\tilde{V}^{(0)i}(k_t, b) = \Xi(k_t) - M^j(k_t)$

*Notation* 2.

$$\tilde{V}^{(1)i}(k_t, b^*) = \min\{\tilde{V}^{(0)i}(k_t, b), \dots$$
$$\max_{b \in \phi^{ex}}[u^i(k_t, b_t) + \delta \textstyle\sum_{k_{t+1} \in K_{t+1}} q(k_{t+1}|k_t, b_t) \max\{M^i(k_{t+1}), \Xi(k_{t+1}) - M^j(k_{t+1})\}]\}$$

Assume that war does not occur at $t$ and let $b^*$ be the arg max with $b_t^*$ the element in the sequence $b^*$ corresponding to time $t$.

**Lemma B.2.** *(Powell or $SC(1)$ sufficient condition for war) If state $k_t$ is reached and $\exists\ i$ such that $M^i(k_t) > \tilde{V}^{(1)i}(k_t, b^*)$, then war occurs at $t$.*

The $SC(1)$ condition essentially combines the $SC(0)$ condition with the insight from Powell (2004). When the min is $\tilde{V}^{(1)i}(k_t, b^*) = \tilde{V}^{(0)i}(k_t, b^*) = \Xi(k_t) - M^j(k_t)$ then, if war results, it arises for the same reasons as in the $SC(0)$ condition. When the min is $u^i(k_t, b_t^*) + \delta \sum_{k_{t+1} \in K_{t+1}} q(k_{t+1}|k_t, , b_t^*) \max\{M^i(k_{t+1}), \Xi(k_{t+1}) - M^j(k_{t+1})\}$, the logic is a little more subtle. Essentially, if this term is the smaller one, then there exists bargains that $j$ prefers to war in period $t$. $b_t^*$ is the bargain that accomplishes this at $t$ while giving $i$ the highest value. Since $j$ has the option of going to war at time $t+1$, $i$ cannot hope to get more than $\Xi(k_{t+1}) - M^j(k_{t+1})$ from a bargain at any particular state $k_{t+1}$. The max term comes from the fact $i$ also has the ability to initiate war in $t+1$, and if condition $SC(0)$ is met at that time, $i$ certainly will.

Consider what would be expected to happen in $t+1$ under rationality when $V^i(k_t,b)$ is known exactly. There will be some states $k_{t+1}$ that fall into the set $K^P_{t+1}(b)$, where $i$ receives some bargain in $k_{t+1}$. In those states, $\tilde{V}^{(1)i}(k_t,b^*)$ assigns a value of $\Xi(k_{t+1}) - M^j(k_{t+1})$ which is the highest present value bargain value $i$ can possibly expect at $k_{t+1}$ without considering any further restrictions. Other states $k_{t+1}$ will fall into $K^W_{t+1}(b)$, in which rationality predicts war. $SC(1)$ uses the war value only in the subset of those states in which $M^i(k_{t+1}) > \Xi(k_{t+1}) - M^j(k_{t+1})$. When this is not true, $SC(1)$ assigns the value $\Xi(k_{t+1}) - M^j(k_{t+1})$ which is a higher value than achieved under rationality. Since $u(k_t,b^*_t)$ is the highest possible bargain attainable at time $t$, this means that $\tilde{V}^{(1)i}(k_t,b^*) \geq V^i(k_t,b)$ for all possible bargains. Hence, $SC(1)$ is a sufficient condition. Since one only need consider $u^i(k_t,b^*_t) + \delta \sum_{k_{t+1} \in K_{t+1}} q(k_{t+1}|k_t) \max\{M^i(k_{t+1}), \Xi(k_{t+1}) - M^j(k_{t+1})\}$ when it is in fact smaller than the condition in $SC(0)$ (which is why we have the minimization term), this is a broader condition than $SC(0)$.

One can go further by building up to a broader $SC(n)$ condition. First note that the inequality in the $SC(1)$ condition defines a set of states $k_t$ in which war definitely occurs. For simplicity, imagine that the period one is in interested in is $t=0$ and that one in fact knows the $SC(n)$ condition for war. Let $k_t \in K^{W(n)i}_t$ be a state in which $i$ declares war in $t$ according to the $SC(n)$ war condition and $K^{W(n)i}_t$ the set of all such states. Let $k^t \in K^{P_t(n)}$ be a sequence of states from 0 to $t$ such that war does not occur according to the $SC(n)$ condition for war and $K^{P_t(n)}$ the set of all such sequences.

*Notation* 3. $\hat{M}^{(n)i}_t(k_0,b^*) \equiv \sum_{k^{t-1} \in K^{P_{t-1}(n)}} \rho(k^{t-1}|k_0,b^*) \sum_{k_t \in K^{W(n)i}_t} q(k_t|k^{t-1}_{t-1},b^*_{t-1}) M^i(k_t)$

This piece of notation deserves some unpacking. First, the bargain $b^*$ is implicit at every point in time given $k$, so one could suppress it and use only $\hat{M}^{(n)i}_t(k_0)$. $\hat{M}^{(n)i}_t(k_0)$ is the expected value of wars initiated by $i$ that will occur at period $t$ according to $SC(n)$. The first summation over $\rho(\cdot)$ gives the probability of arriving at $t-1$ peacefully. The summations over $q(\cdot)$ gives the probabilities of arriving at states in $t$ from states in $t-1$.

I will also need to keep track of $i$'s expected value at states where $j$ initiates a war.

*Notation* 4.

$$\hat{M}_t^{(n)ij}(k_0,b^*) \equiv \sum_{k^{t-1}\in K^{P_{t-1}(n)}} \rho(k^{t-1}|k_0,b^*) \sum_{k_t \in (K_t^{W(n)j} \setminus K_t^{W(n)i})} q(k_t|k_{t-1}^{t-1},b_{t-1}^*)M^i(k_t)$$

This is the same as Notation 3 except for the following modification. I am using the probability of being in a state where $j$ prefers war at some time $t$ and $i$'s value for war (I exclude cases where they both prefer war since I included those in Notation 3 and want to avoid double counting). This is because I am interested in cases where $j$ can calculate he will be forced into war in the future.

I denote the expected value of the utility at $t$ at the maximum bargain for $i$ under the $SC(n)$ as

*Notation* 5. $\mu_t^{(n)i}(k_0,b^*) \equiv \sum_{k^t\in K^{P_t(n)}} \rho(k^t|k_0,b^*)u^i(k_t,b_t^*)$

It is also convenient to define notation for the maximum between the war value and $\tilde{V}^{(1)i}(k_t,b^*)$. Since $\tilde{V}^{(1)i}(k_t,b^*)$ is a bound on $V^i(k_t,b)$ and war is always possible, the max of these terms together always bounds above the true value function of a given state.

*Notation* 6. $\Upsilon^i(k_t) \equiv \max\{M^i(k_t),\tilde{V}^{(1)i}(k_t,b^*)\}$

Let the expected value at $t=0$ of at some future date be denoted:

*Notation* 7. $\hat{\Upsilon}_t^i(k_0) = \sum_{k^t\in K^{P_t(n)}} \rho(k^t|k_0,b^*)\Upsilon^i(k_t)$

Using this notation we can now state the $SC(2)$ condition.

**Lemma B.3.** *($SC(2)$ sufficient condition for war) If state $k_t$ is reached and $\exists\ i$ such that*

$$M^i(k_0) > \min\{\tilde{V}^{(1)i}(k_t, b^*), \dots$$
$$u^i(k_0, b_0^*) + \delta[\mu_1^{(1)i}(k_0, b^*) + \hat{M}_1^{(1)i}(k_0, b^*) + \hat{M}_1^{(1)ij}(k_0, b^*)] + \delta^2 \hat{\Upsilon}_2^i(k_0)\}$$

*then war occurs at $t$.*

One can now derive broader sufficient conditions by iterating the logic of the $SC(2)$ sufficient condition further into future. For $SC(n)$, the RHS of the condition is denoted as follows:

*Notation* 8.

$$\tilde{V}^{(n)i}(k_0, b^*) \equiv \min\{\tilde{V}^{(n-1)i}(k_t, b^*), u^i(k_0, b_0^*) +$$
$$+ \sum_{t=1}^{n-1} \delta^t [\mu_t^{(n-1)i}(k_0) + \hat{M}_t^{(n-1)i}(k_0, b_0^*) + \hat{M}_t^{(n-1)ij}(k_0, b_0^*)] + \delta^n \hat{\Upsilon}_n^i(k_0)\}$$

Now I can state the first main result.

**Proposition B.1.** *($SC(n)$ sufficient condition for war) $i$ declares war at $t = 0$ if*

$$M^i(k_0) > \tilde{V}^{(n)i}(k_0, b^*) \tag{B.1}$$

The broadest sufficient condition so far occurs when one considers $SC(n = \infty)$. This condition may be intractable to calculate for some modeling purposes. Thus, using a smaller finite value for $n$ may be necessary. Furthermore, allowing for a general value of $n$ allows one to investigate some types of bounded rationality in this setting. Lower $n$ values correspond to higher degrees of myopia in the decision makers. Also, this condition can be modified to allow for different degrees of bounded rationality amongst the players by using different values of $n$ for $i$ and $j$. One can also directly impose bounded rationality by restricting the set of available bargains $\phi^{ex}$.

Using Proposition B.1, one knows that there are certain values of $k_t$ where, only looking at the condition for $i$, there does not exist a bargain, $b$, such that war does not occur. In constructing $\tilde{V}^{(n)i}(k_t, b^*)$, I considered wars in the future that $j$ might cause, but I did so while letting $j$ have the best possible bargain it could hope for without regard to the bargain I was assuming for $i$. This then is also the case for inequality (B.1) itself. $i$ may avoid war by considering bargain $b^{(i)*}$ and $j$ may avoid it with $b^{(j)*}$, but this does not mean that there exists a mutually compatible bargain that prevents war.

If I can calculate $\tilde{V}^{(n)i}(k_t, b)$ for all possible $b \in \phi^{ex}$, I can find the wars that occur when there does not exist a mutually compatible $b$. The rest of this section will proceed to do this and build to the broadest sufficient condition, which is shown in Proposition B.3. However, it should be noted that Proposition B.1 (as well as Proposition B.2 which is proved as a step towards Proposition B.3) can be a great deal more tractable to calculate than Proposition B.3. This is due to the fact that Proposition B.3 requires one to calculate values under every possible bargain $b$ as opposed to just the optimal bargain $b^*$. Since bargains $b$ assign a value for $i$ and $j$ in every state $k_t$, this can cause tractability issues for certain specifications. Modelers can choose from the menu of sufficient conditions in order to find the appropriate trade-off between tractability and precision for their model.

Consider the value $\tilde{V}^{(n)i}(k_t, b)$ where $b \in \phi^{ex}$. Define the subset $\phi^{(i)en}(k_t, n) \subseteq \phi^{ex}$ where $\phi^{(i)en}(k_t, n)$ is the set of bargains so that $M^i(k_t) \leq \tilde{V}^{(n)i}(k_t, b) \; \forall \; b \in \phi^{(i)en}(k_t)$. This means that $\phi^{(i)en}(k_t, n)$ is the set of bargains that are potentially peaceful under Proposition B.1 for $i$ at $k_t$. If $SC(n)$ is satisfied, the $\phi^{(i)en}(k_t, n)$ is the empty set. There are no bargains for which $i$ will put off war. One can define a similar set for $j$. Let the intersection of these two sets be $\phi^{en}(k_t, n) \equiv \phi^{(i)en}(k_t, n) \cap \phi^{(j)en}(k_t, n)$ so that $\phi^{en}(k_t, n)$ is the set of possibly peaceful bargains that are endogenously restricted by applying $SC(n)$ to both $i$ and $j$ at $k_t$.

Leveraging this endogenous restriction on bargains gives us a new sufficient condition for war.

**Proposition B.2.** *$\phi^{en}(k_t,n)$ has the following three characteristics.*

*1. ($SC(n,1)$ sufficient condition for war) If $\phi^{en}(k_t,n)=\emptyset$, then war occurs at $k_t$.*

*2. If $\phi^{en}(k_t,n)\neq\emptyset$, then for any peaceful bargain $b$ at $k_t$ we have that $b\in\phi^{en}(k_t,n)$.*

*3. If war occurs under $SC(n)$, then it occurs under $SC(n,1)$. The converse does not hold.*

Now that we have the sufficient condition $SC(n,1)$, it is natural to push the insight further by applying it to states as far in the future as possible. Define $\phi_t^{en}(k_t,n)$ as the set of bargains $b_t$ such that $b_t$ is the bargain at $t$ in some sequence $b\in\phi^{en}(k_t,n)$. I now introduce notation necessary to build $SC(n,2)$. First, let $K_t^{W(n,1)i}(b)$ be the set of $k_t$ that result in war when the bargain is $b$ and the condition for war is $SC(n,1)$ and let $K^{P_t(n,1)}(b)$ be the set of sequences that are peaceful under bargain $b$ and condition $SC(n,1)$. I keep track of the war values using the following notation.

*Notation* 9.

$$W_t^{(n,1)i}(k_0,b)\equiv$$

$$\textstyle\sum_{k^{t-1}\in K^{P_{t-1}(n,1)}(b)}\rho(k^{t-1}|k_0,b)\sum_{k_t\in(K_t^{W(n,1)i}(b)\cup K_t^{W(n,1)j}(b))}q(k_t|k_{t-1}^{t-1},b_{t-1})M^i(k_t)$$

This simply assigns a war value in every state in which war would occur under $SC(n,1)$ for a given bargain $b$. Next, I provide notation for when the bargain in question leads to a peaceful outcome in a given state.

*Notation* 10. $v_t^{(n,1)i}(k_0,b)\equiv\sum_{k^t\in K^{P_t(n,1)}(b)}\rho(k^t|k_0,b)u^i(k_t,b_t)$

Notation 10 is the instantaneous value of a bargain $b$ in given peaceful state $k_t$.

*Notation* 11. $P_t^{(n,1)i}(k_0,b) \equiv \sum_{k^t \in K^{P_t(n,1)}(b)} \rho(k^t|k_0,b)\tilde{V}^{(n)i}(k_t,b)$

Notation 11 provides the continuation value of a bargain $b$ in a given peaceful state $k_t$. Next I define the following value function:

*Notation* 12. $\Lambda^{(n,2)i}(k_0,b) \equiv v_0^{(n,1)i}(k_0,b) + \delta[P_1^{(n,1)i}(k_0,b) + W_1^{(n,1)i}(k_0,b)]$

Finally, this allows us to define the next level of endogenously restricted bargains.

**Definition B.1.** $b \in \phi^{en}(k_t,n,2)$ iff $M^i(k_t) \leq \Lambda^{(n,2)i}(k_t,b)$ for all $i$.

The following lemma shows uses this definition to show $SC(n,2)$. For simplicity we start in state $k_0$.

**Lemma B.4.** $\phi^{en}(k_t,n,2)$ has the following three characteristics.

*1. ($SC(n,2)$ sufficient condition for war) If $\phi^{en}(k_0,n,2) = \emptyset$, then war occurs at $k_t$.*

*2. If $\phi^{en}(k_t,n,2) \neq \emptyset$, then for any peaceful bargain b at $k_t$ we have that $b \in \phi^{en}(k_t,n,2)$.*

*3. If war occurs under $SC(n,1)$, then it occurs under $SC(n,2)$. The converse does not hold.*

In order to state Proposition B.3, generalize Notations 9-11 with the generic natural number $m$ replacing 1. Then one can replace Notation 12 with the following:

*Notation* 13.

$$\Lambda^{(n,m)i}(k_0,b) \equiv$$
$$\sum_{t=1}^{m-1} \delta^t [v_{t-1}^{(n,m-1)i}(k_0,b) + W_t^{(n,m-1)i}(k_0,b)] + \delta^{m-1} P_{m-1}^{(n,m-1)i}(k_0,b)$$

Then this allows us to define are general endogenous constraint on bargaining.

**Definition B.2.** $b \in \phi^{en}(k_t,n,m)$ iff $M^i(k_t) \leq \Lambda^{(n,m)i}(k_t,b)$ for all $i$.

Now we finally state our main result and broadest sufficient condition for war.

**Proposition B.3.** $\phi^{en}(k_t, n, m)$ has the following three characteristics.

1. *($SC(n,m)$ sufficient condition for war) If $\phi^{en}(k_0, n, m) = \emptyset$, then war occurs at $k_t$.*
2. *If $\phi^{en}(k_t, n, m) \neq \emptyset$, then for any peaceful bargain $b$ at $k_t$ we have that $b \in \phi^{en}(k_t, n, m)$.*
3. *If war occurs under $SC(n, m-1)$, then it occurs under $SC(n,m)$.*

Statement 1 ensures that we have indeed found a sufficient condition for war. Statement 2 allows us to conclude that any peaceful bargain that may be found must be part of the set $\phi^{en}(k_t, n, m)$. Statement 3 allows us to conclude that $SC(n,m)$ is the broadest type of sufficient condition found here. Of course, if it is tractable to calculate, $SC(\infty, \infty)$ would be the broadest condition of this type.

# C Proofs

## C.1 Chapter 1 Proofs

**Proof of Lemma 1.1:** Consider the subgame beginning in period $T$. The action for $S$ at $T$ is to offer a bargain value $x_T$ and then for $R$ to play $a$ for accepting the bargain or $w$ for war contingent on $x_T$. Consider an $x_T^{'}$ such that $1-x_T^{'} \geq (1-\delta)M_{RT}$. This implies a payoff in period $T$ of at most $(1-\delta)M_{RT}$ for $R$. If $R's$ strategy is $w$ for $x_T^{'}$, then $R$ has a positive deviation to play $a$ this period even when receiving its worst possible value next period of $M_{RT}$, since $1-x_T^{'}+\delta M_{RT} \geq M_{RT}$.[70] Therefore, $R$ must play $a$ for $x_T^{'}$. Since $R$ must accept all values $1-x_T \geq (1-\delta)M_{RT}$, $S$ must not propose $1-x_T > (1-\delta)M_{RT}$, since any proposal such that $1-x_T > (1-\delta)M_{RT}$ has a profitable deviation to $1-x_T^{'} = (1-\delta)M_{RT}$. By this logic, $1-x_t \leq (1-\delta)M_{RT}$ for all $t \geq T$. This means the most optimistic $R$ can be of its value for future periods is that $1-x_t = (1-\delta)M_{RT}$ for all $t \geq T$. This most optimistic assessment implies that $R$'s valuation of peace starting in $T+1$ cannot be greater than $M_{RT}$. This means that for $1-x_T^{''} < (1-\delta)M_{RT}$, $R$ has profitable deviation from the action $a$ to $w$, since $M_{RT} > 1-x_T^{''}+\delta M_{RT}$. Therefore, in any period $t \geq T$, $R$'s strategy is $a$ if $1-x_t \geq (1-\delta)M_{RT}$ and $w$ if $x_t < (1-\delta)M_{RT}$. This in turn implies that $S$ must play $(x_t, A_{St}) = (1-(1-\delta)M_{RT}, a)$ in all $t \geq T$. This gives utility values $(U_{ST}, U_{RT}) = (B-M_{RT}, M_{RT})$ for the game starting in $T$.

(Sufficiency): $M_{S0} > B-\delta^T M_{RT} \Rightarrow$ War occurs in any SPE of the game, $\Gamma$, under bargaining protocol $\mathbb{S}$ at $0 \leq t < T$.

Let $U_{it}$ represent the total utility received by $i$ starting at $t$ if $\Gamma$ is peaceful. If $S$ receives the entire bargain value in all $t < T$, then $U_{S0} = \sum_{t=0}^{T-1} \delta^t + \delta^T (B-M_{RT}) = B-\delta^T M_{RT}$. Therefore, if $M_{S0} > B-\delta^T M_{RT}$, then war at $t=0$ is preferred to any peaceful bargain by $S$.

[70] Remember, I assume peace in case where war and peace give equal value. This can be thought of as an $\varepsilon > 0$ value on peace where $\varepsilon$ is small.

Therefore, if war does not occur at $t = 0$, it must be that war occurs at $0 < t < T$, otherwise $S$ has a positive deviation to $w$ at $t = 0$. QED.

**Proof of Lemma 1.2:** First, the converse statement. By definition, $\mathbb{S}$ assigns a value $x_t \in [0,1]$ for all $t$ so therefore $\exists\ P \in \mathbb{P}$ such that $\mathbb{S}$ corresponds to $P$. Therefore, the converse statement holds trivially.

Now, assume that $w$ is played at time $\tau$ under $S$, but that $\exists\ P^{'} \in P$ such that $w$ is not played at some $0 \leq t \leq \tau$ in $P^{'}$. Then it must be that $w$ is not played at $\tau$. Since $w$ was not played before $\tau$, it must be that either $w$ is never played or that $w$ is played at $t^{'} > \tau$.

In the first case, this means that $\sum_{\tau}^{\infty} \delta^{t-\tau} x_t(P^{'}) \geq M_{S\tau}$. But since $w$ is played at $\tau$ under $\mathbb{S}$, there must not be a positive deviation to $a$ under $\mathbb{S}$, so that $M_{S\tau} > \sum_{\tau}^{\infty} \delta^{t-\tau} x_t(\mathbb{S}^{'})$ where $\mathbb{S}^{'}$ generates the largest peaceful bargain value for $S$ that is credible. This implies that $\sum_{\tau}^{\infty} \delta^{t-\tau} x_t(P^{'}) > \sum_{\tau}^{\infty} \delta^{t-\tau} x_t(\mathbb{S}^{'})$. Since $S$ makes the offers of $x_t$ under $\mathbb{S}$, it must be that $P^{'}$ is an incredible sequence of offers for $S$. $P^{'}$ can be incredible for $S$ only because it is not minmax compatible, in which case there is war at some time $t$, or because $S$ cannot commit to the sequence $P^{'}$ because of its offer power. This means that there is a $\tau^{'} \geq \tau$ such that $\sum_{\tau^{'}}^{\infty} \delta^{t-\tau^{'}} x_t(\mathbb{S}^{'}) > \sum_{\tau^{'}}^{\infty} \delta^{t-\tau^{'}} x_t(P^{'})$. Therefore, $\sum_{\tau}^{\tau^{'}-1} \delta^{t-\tau} x_t(P^{'}) > \sum_{\tau}^{\tau^{'}-1} \delta^{t-\tau} x_t(\mathbb{S}^{'})$. But since $S$ makes the offers in $\mathbb{S}^{'}$, $S$ can choose $\{x_{\tau}, ..., x_{\tau^{'}}\} = \{x_{\tau}(P^{'}), ..., x_{\tau^{'}}(P^{'})\}$ unless this is not minmax compatible for $R$. By definition of $\theta$, from the perspective of time $\tau$, $M_{Rt}$ is increasing in present value terms over the range $\tau$ to $\tau^{'}$ until $\tau^{''}$ such that $\tau \leq \tau^{''} \leq \tau^{'}$, at which point it is constantly weakly decreasing in present value terms.[71] If $t$ is in the range in which $M_{Rt}$ is increasing in present value terms, $R$ will accept any bargain in period $t$ since $\delta M_{Rt+1} \geq M_{Rt}$. Therefore if $\tau^{''} = \tau^{'}$, then $\{x_{\tau}, ..., x_{\tau^{'}}\} = \{x_{\tau}(P^{'}), ..., x_{\tau^{'}}(P^{'})\}$ is in fact minmax compatible for $R$ and there exists a sequence such that $\sum_{\tau}^{\tau^{'}-1} \delta^{t-\tau} x_t(\mathbb{S}^{'}) \geq$

[71]It is constantly weakly decreasing if $\tau \geq T$ or $\theta \leq \frac{1}{\delta}$. It is constantly increasing until $\tau^{''}$ if $\theta > \frac{1}{\delta}$ and $T = \tau^{''}$ or $\tau^{'} = \tau^{''}$.

$\sum_{\tau}^{\tau'-1} \delta^{t-\tau} x_t(P')$ implies $\sum_{\tau}^{\infty} \delta^{t-\tau} x_t(\mathbb{S}') > \sum_{\tau}^{\infty} \delta^{t-\tau} x_t(P')$ a contradiction. If $\tau'' < \tau'$, then $\{x_{\tau''}, ..., x_{\tau'}\} = \{x_{\tau''}(P'), ..., x_{\tau'}(P')\}$ is incredible under $\mathbb{S}$. Since $R$ does not play $w$ at $\tau''$ under $P'$ but it is minmax incompatible for $R$ under $\mathbb{S}$, it must be that $\sum_{\tau''}^{\infty} \delta^{t-\tau'} \left(1 - x_t(P')\right) \geq M_{R\tau''} > \sum_{\tau''}^{\infty} \delta^{t-\tau} \left(1 - x_t(\mathbb{S}')\right)$ which in turn implies that $\sum_{\tau''}^{\infty} \delta^{t-\tau'} x_t(\mathbb{S}') \geq \sum_{\tau''}^{\infty} \delta^{t-\tau'} x_t(P')$. But, in all periods $t < \tau''$, $\mathbb{S}'$ can always at least mimic the $P'$ offer without causing war, which implies that $\sum_{\tau}^{\infty} \delta^{t-\tau} x_t(\mathbb{S}') \geq \sum_{\tau}^{\infty} \delta^{t-\tau} x_t(P')$, which contradicts the original premise.

Now assume that $w$ is played at $t' > \tau$. Since $M_{St}$ is necessarily decreasing in present value terms by definition of $\theta$, this implies that $\sum_{\tau}^{t'-1} \delta^{t-\tau} x_t(P') > \sum_{\tau}^{t'-1} \delta^{t-\tau} x_t(\mathbb{S}')$. If $R$ prefers this bargain to waiting, then it must be that $\delta M_{Rt+1} \geq M_{Rt}$ holds in this range and under $\mathbb{S}$ any bargain is possible, therefore $\sum_{\tau}^{t'-1} \delta^{t-\tau} x_t(\mathbb{S}') \geq \sum_{\tau}^{t'-1} \delta^{t-\tau} x_t(P')$ a contradiction. Otherwise, $R$ prefers war at some $t'' < t'$. But by the same logic, if $R$ prefers war at $t''$ to war at $\tau$, then any bargain is possible under $\mathbb{S}$ in that range, therefore it must be that if $t'' > \tau$ then $\sum_{\tau}^{t''-1} \delta^{t-\tau} x_t(\mathbb{S}') \geq \sum_{\tau}^{t''-1} \delta^{t-\tau} x_t(P')$, which again contradicts the premise that $P'$ can delay war when $\mathbb{S}'$ does not. If $t'' < \tau$, then war may occur earlier under $P'$, but not later, again contradicting that $w$ is played at some $t' > \tau$. This exhausts all cases. QED.

**Proof of Proposition 1.2:** By Lemma 1.1, if (1.3) holds, then $w$ is played at some time $\tau'$ where $0 < \tau' < T$. By Lemma 1.2, $w$ is then played at some $t$ such that $0 \leq t \leq \tau'$ in $\Gamma$ for all protocols $P \in \mathbb{P}$. Let $\tau = t$. QED.

**Proof of Corollary 1.1:** Plugging in for inequality (1.3) gives

$$\frac{(1-p_0)(1-d)}{1-\delta} > \frac{1}{1-\delta} - \frac{(\delta\theta)^T p_0(1-d)}{1-\delta}$$
$$(1-p_0)(1-d) > 1-(\delta\theta)^T p_0(1-d)$$
$$\theta^T > \frac{1-(1-p_0)(1-d)}{\delta^T p_0(1-d)}$$
$$\theta > \frac{1}{\delta}\left(1+\frac{d}{p_0(1-d)}\right)^{\frac{1}{T}}$$

The terms within the parenthesis add to be greater than 1. Therefore, the RHS is decreasing in $T$. Therefore, if the inequality holds for $T=1$, then it holds for all $T \geq 1$. The converse does not hold by the counterfactual analyzed in Chapter 1 where $\delta = 0.96$, $d = 0.2$, and $p_0 = \frac{1}{3}$. QED.

**Proof of Proposition 1.3:** Assume not so that $\theta \leq \frac{1}{\delta}$, $P^{'}$ is minmax compatible at $t=0$, but war occurs at $t=0$. Since $P^{'}$ is minmax compatible at $t=0$, both $S$ and $R$ prefer $P^{'}$ to war at $t=0$ if accepting the bargain in all periods $t>0$ is credible. Therefore, accepting the bargain implied by $P^{'}$ in all periods $t>0$ must be incredible, therefore war occurs at some point $t^{'}>0$ if the bargain is not accepted today. $S$ and $R$ choose not to wait for war until $t^{'}$ only if $M_{S0} > \sum_{t=0}^{t^{'}-1}\delta^t x_t(P^{'}) + \delta^{t^{'}} M_{St^{'}}$ or $M_{R0} > \sum_{t=0}^{t^{'}-1}\delta^t\left(1-x_t(P^{'})\right) + \delta^{t^{'}} M_{Rt^{'}}$. This gives three cases. First imagine that it is the case that $M_{S0} > \sum_{t=0}^{t^{'}-1}\delta^t x_t(P^{'}) + \delta^{t^{'}} M_{St^{'}}$ and $M_{R0} > \sum_{t=0}^{t^{'}-1}\delta^t\left(1-x_t(P^{'})\right) + \delta^{t^{'}} M_{Rt^{'}}$. Adding gives

$$M_{S0}+M_{R0} > \sum_{t=0}^{t^{'}-1}\delta^t + \delta^{t^{'}}\left(M_{St^{'}}+M_{Rt^{'}}\right)$$

Using constant inefficiency of war gives

$$B(1-d) > \sum_{t=0}^{t'-1} \delta^t + \delta^{t'} \left(B(1-d)\right)$$
$$B - dB > B - \delta^{t'} dB$$

which is false since $\delta < 1$. Now check the case where $M_{R0} > \sum_{t=0}^{t'-1} \delta^t \left(1 - x_t(P')\right) + \delta^{t'} M_{Rt'}$, but $M_{S0} \leq \sum_{t=0}^{t'-1} \delta^t x_t(P') + \delta^{t'} M_{St'}$. By $\theta \leq \frac{1}{\delta}$, $\delta^{t'} M_{Rt'} \leq M_{R0}$, therefore the first inequality gives $\sum_{t=0}^{t'-1} \delta^t \left(1 - x_t(P')\right) < 0$ a contradiction since $x_t(P') \leq 1$. For the last case where $M_{S0} > \sum_{t=0}^{t'-1} \delta^t x_t(P') + \delta^{t'} M_{St'}$, but $M_{R0} \leq \sum_{t=0}^{t'-1} \delta^t \left(1 - x_t(P')\right) + \delta^{t'} M_{Rt'}$, there are two subcases. In the first subcase, $P'$ is not minmax compatible for $S$ at $t'$. Since $M_{St'} > \sum_{t=t'}^{\infty} \delta^{t-t'} x_t(P')$, then $\sum_{t=0}^{t'-1} \delta^t x_t(P') + \delta^{t'} M_{St'} > \sum_{t=0}^{\infty} \delta^t x_t(P')$. But since $P'$ is minmax compatible at $t = 0$, $\sum_{t=0}^{\infty} \delta^t x_t(P') \geq M_{S0}$ implies $\sum_{t=0}^{t'-1} \delta^t x_t(P') + \delta^{t'} M_{St'} > M_{S0}$ a contradiction. The other subcase implies that $P'$ is minmax compatible for $S$, therefore it must not be for $R$ since war occurs in the future, but I have assumed that the bargain is minmax compatible for $R$ at all $t$ a contradiction.

This exhausts all cases. QED.

**Proof of Proposition 1.4:** By Proposition 1.2, war must occur at some $\tau$ such that $0 \leq \tau < T$.

Assume that period $t$ is reached and that (1.4) is satisfied. Assume the contrary and that $S$ waits until some $t < t' < T$ to play $w$. First, assume that $t' = t + 1$. For this to be a positive deviation it must be that

$$1+\delta\frac{(1-\theta p_t)(1-d)}{1-\delta} \geq \frac{(1-p_t)(1-d)}{1-\delta}$$
$$\frac{1-\delta}{(1-d)}+\delta-\delta\theta p_t \geq 1-p_t$$
$$\delta\theta p_t \leq \frac{1-\delta}{(1-d)}+\delta+p_t-1$$
$$\theta \leq \frac{1}{\delta}+\frac{1}{p_t}+\frac{1}{\delta p_t}\left[\frac{1-\delta}{1-d}-1\right]$$
$$\theta \leq \frac{1}{\delta}+\frac{1}{\delta p_t}\left[\frac{\delta-1-\delta d+d}{1-d}+\frac{1-\delta}{1-d}\right]$$
$$\theta \leq \frac{1}{\delta}+\frac{1}{p_t}\left[\frac{d(1-\delta)}{\delta(1-d)}\right]$$

But this contradicts (1.4) exactly. Now consider $t^{'} > t+1$. Define $s = t^{'}-1$. Since $\theta > 1$,

$$\frac{1}{\delta}+\frac{1}{p_{t'}}\left[\frac{d(1-\delta)}{\delta(1-d)}\right] > \frac{1}{\delta}+\frac{1}{p_s}\left[\frac{d(1-\delta)}{\delta(1-d)}\right]$$

Therefore, war is preferred by $S$ at $s$ to war at $t^{'}$. Since this is true for all $t^{'} > t$, war must occur at $t$. QED.

**Proof of Proposition 1.5:** Since (1.4) holds, $\theta > \frac{1}{\delta}$, therefore, for $T$ large enough, (1.3) is satisfied. Since (1.3) and (1.4) are satisfied, sufficiency follows from Proposition 1.4. In order to show necessity, assume that (1.4) does not hold, yet either $S$ or $R$ chooses $w$ at $t=0$. Since $\delta M_{R1} > M_{R0}$ by virtue of $\theta > \frac{1}{\delta}$, $R$ has a positive deviation to $a$ for any $x_0 \in [0,1]$. Since $R$ will accept any offer, $S$ prefers $w$ to $a$ only if

$$\frac{(1-p_0)(1-d)}{1-\delta} > 1+\delta\frac{(1-\theta p_0)(1-d)}{1-\delta}$$

But from Proposition 1.4, this holds exactly when (1.4) holds. QED.

**Proof of Proposition 1.6:** If (1.5) $\iff$ (1.3), Proposition1. 6 follows from Proposition 1.5. (1.3) states that

$$M_{S0} > B - \delta^T M_{RT}$$

Plugging in gives

$$\begin{aligned}
\tfrac{(1-p_0)(1-d)}{1-\delta} &> \tfrac{1}{1-\delta} - \tfrac{\delta^T \theta^T p_0(1-d)}{1-\delta} \\
1 - p_0 &> \tfrac{1}{1-d} - \delta^T \theta^T p_0 \\
\theta &> \left[\tfrac{1}{\delta^T p_0}\left(\tfrac{1}{1-d} + p_0 - 1\right)\right]^{\frac{1}{T}} \\
\theta &> \tfrac{1}{\delta}\left(1 + \tfrac{d}{p_0(1-d)}\right)^{\frac{1}{T}}
\end{aligned}$$

which is (1.5), therefore $(1.5) \iff (1.3)$. QED.

## C.2 Chapter 2 Proofs

### C.2.1 Formally defining the game

Consider two games – one with commitment and one without – with three players $H$, $F$, and $L$. In the game without commitment, players' action sets are

$$\mathscr{A}_H = \{B | B \geq 0\} \quad \mathscr{A}_F = \{x_t | x_t \in [0,1]\} \quad \mathscr{A}_F = \{\text{accept, reject}\}$$

The stage game timing is:

1. Players observe $y_t$.
2. $H$ chooses $B_t \in \mathscr{A}_H$
3. $F$ chooses $x_t \in \mathscr{A}_F$

4. $L$ chooses $\alpha_t \in \mathscr{A}_L$.

Let $a_t = (B_t, x_t, \alpha_t)$ be an action profile chosen when the status quo is $y_t$. Stage game payoffs are as follows:

$$u_H(a_t; y_t) = \begin{cases} x_t & \text{if } \alpha_t = \text{accept} \\ M_H(y_t) & \text{if } \alpha_t = \text{reject} \end{cases}$$

$$u_F(a_t; y_t) = \begin{cases} 1 - x_t & \text{if } \alpha_t = \text{accept} \\ M_F(y_t) & \text{if } \alpha_t = \text{reject} \end{cases}$$

$$u_L(a_t; y_t) = \begin{cases} x_t & \text{if } \alpha_t = \text{accept} \\ M_L(y_t, B_t) & \text{if } \alpha_t = \text{reject} \end{cases}$$

We denote a strategy for player $i$ in the dynamic game by $\sigma_i$. Because the solution concept we focus on is Markov perfect equilibrium, we forgo a formal definition of the strategy space, focusing instead on strategies that condition only on the current status quo and on the actions of other players that can be observed in the stage game. Thus $y \mapsto \sigma_H(y) \in \mathscr{A}_H$, $(y,B) \mapsto \sigma_F(y,B) \in \mathscr{A}_F$, and $(y,B,x) \mapsto \sigma_L(y,B,x) \in \mathscr{A}_L$.

In the game with commitment, we must be careful to allow $H$ to choose only once. Let $\theta = 0$ if it is the first period of the game and let $\theta = 1$ otherwise. Define

$$\mathscr{A}_H^0 = \{B | B \geq 0\} \quad \text{and} \quad \mathscr{A}_H^1 = \{B_0\}.$$

$H$'s strategies now have the form $(\theta, y) \mapsto \sigma_H(\theta, y) \in \mathscr{A}_H^\theta$. The strategies of $F$ and $L$ remain unchanged.

Dynamic payoffs are represented by functions $(\cdot, y) \mapsto V_i(\cdot, y)$ for each $i$. We allow payoffs to be functions of both a strategy profile or the sequence of offers and biases implied by a strategy profile. Thus if $\sigma = (\sigma_H, \sigma_F, \sigma_L)$ implies $(\mathbf{x}, \mathbf{B}), = ((x_0, x_1, \ldots), (B_0, B_1, \ldots))$, we have $V_i(\sigma, y) = V_i((\mathbf{x}, \mathbf{B}), y)$. If the game is with commitment, we will frequently drop the implied sequence of leader biases and write instead $V_i(\sigma, y, B)$ or $V_i(\mathbf{x}, y, B)$ for the associated payoffs. Finally, let $v_i(\cdot, y) = (1-\delta_i)V_i(\cdot, y)$.

## Proofs

### Proof of Lemma 2.1

To see that $M_L(y, B)$ is concave if $B > 1$, note that

$$\begin{aligned}(1-\delta)M_L(y,B) &= [(1-C)y + BG(1-y)]\, p(y) + [(1-C-G)y]\,(1-p(y)) \\ &= (1-C-G)y + G\,(B(1-y)+y)\, p(y).\end{aligned}$$

Then,

$$\frac{\partial^2}{\partial y^2}(1-\delta)M_L(y,B) = G\,(B(1-y)+y)\, p''(y) - 2G(B-1)p'(y) < 0.$$

Existence of $B^\dagger(y)$ follows by inspection. In fact,

$$B^\dagger(y) = \frac{C}{Gp(y)(1-y)} + 1.$$

It can be shown that $B^\dagger(y)$ is strictly convex; thus there is a unique $B^\dagger$ solving $\min_{y \in [0,1]} B^\dagger(y)$. That $B^\dagger > 1$ follows from the definition of $B^\dagger(y)$. QED.

**Proof of Proposition 2.1**

**Definition** (Definition of equilibrium strategy profile for Proposition 2.1)**.** Let $h(x,B) = (1-\delta)x + \delta m_L(x,B)$ and if $\exists\, x$ such that $h(x,B) = m_L(y,B)$ let

$$\chi(y,B) = \min\{x \in [0,1] | h(x,B) = m_L(y,B)\}\};$$

otherwise let $\chi(y,B) = 0$. Define $\sigma^* = (\sigma_F^*, \sigma_L^*)$ where

$$\sigma_F^*(y,B) = \begin{cases} \chi(y,B) & \text{if } m_F(y) + h(\chi(y,B),B) \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\sigma_L^*(y,B) = \begin{cases} \text{accept} & h(x,B) \geq m_L(y,B) \\ \text{reject} & h(x,B) < m_L(y,B). \end{cases}$$

The following notation will be helpful. Let $\mathbf{x}(y) = (x_0(y), x_1(y), \ldots)$ be the sequence of offers defined by $\sigma^*$. Let $\mathscr{W}^0(B) = \{y | M_L(y,B) + M_F(y) > \frac{1}{1-\delta}\}$. Let $\mathscr{W}^n(B) = \{y | x_{n-1}(y) \in \mathscr{W}^0(B)\}$. Let $\mathscr{P}(B) = [0,1] \setminus \cup_{n=0}^{\infty} \mathscr{W}^n(B)$. Note that $y \in \mathscr{W}^0(B)$ implies that war occurs immediately under $\sigma^*$, that $y \in \mathscr{W}^n(B)$ for $n > 0$ implies that $\mathbf{x}(y) = (x_0(y), \ldots, x_{n-1}(y))$, and that $y \in \mathscr{P}(B)$ implies that $\mathbf{x}(y)$ is an infinite sequence such that $V_L(\mathbf{x}(y); y, B) = M_L(y,B)$.

**Lemma C.1.** *$F$ cannot improve on $\sigma_F^*$ if $y \in \mathscr{P}(B)$.*

Note that (1) $F$'s payoff, conditional on peace, is decreasing in $L$'s payoff and (2) for peace to occur, $F$ must offer $L$ at least $M_L(y,B)$. Therefore $\sigma_F^*$ minimizes $L$'s payoff amongst all possible peaceful strategy profiles. Deviations to strategies $\sigma_F'$ which lead to war in finite time decrease the size of the total pie while necessarily still satisfying

$V_L(\sigma_F', \sigma_L^*; y, B) \geq M_L(y, B)$. This implies that $V_F(\sigma_F', \sigma_L^*; y, B) < V_F(\sigma^* | y, B)$. QED.

Suppose that $y \in \mathscr{W}^n(B)$: $F$ has three possible types of deviations. In the first, $F$ might prefer *preemptive war* meaning that $F$ prefers to make an unreasonable offer today, forcing war, to waiting $n$ periods before war occurs. In the second, $F$ might prefer to *demand peace* by offering $x \in P(B)$ such that $x < \chi(y, B)$, grabbing more for themselves today while ensuring peace in the future. In the third, $F$ might prefer to *buy peace* by offering $x \in \mathscr{P}(B)$ such $x > \chi(y, B)$ jumping over future war sets.

**Lemma C.2.** *$F$ has no incentive to make preemptive war if $y \in \mathscr{W}^n(B)$ for $n \geq 1$.*

Suppose $y \in \mathscr{W}^n(B)$ for $n \geq 1$ and suppose that $F$ has the incentive to preempt. Let $x_0, x_1, \ldots, x_n$ be the sequence of offers made by $F$ under $\sigma^*$. Then the following are true:

$$x_t + \delta M_L(x_t, B) = M_L(x_{t-1}, B) \quad \forall t \in \{0, \ldots, n-1\}, \tag{C.1}$$

$$M_F(x_t) + M_L(x_t, B) \leq \frac{1}{1-\delta} \quad \forall t \in \{0, \ldots, n-1\}, \quad \text{and} \tag{C.2}$$

$$M_F(x_n) + M_L(x_n, B) > \frac{1}{1-\delta}. \tag{C.3}$$

Since $F$ has the incentive to preempt,

$$M_F(y) > \sum_{t=0}^{n} \delta^t (1 - x_t) + \delta^{n+1} M_F(x_n, B).$$

By inequality (C.3), this implies

$$\begin{aligned} M_F(y) &> \sum_{t=0}^{n} \delta^t (1 - x_t) + \delta^{n+1} \left( \frac{1}{1-\delta} - M_L(x_n, B) \right) \\ &= \frac{1}{1-\delta} - \left( \sum_{t=0}^{n} \delta^t x_t + \delta^{n+1} M_L(x_n, B) \right). \end{aligned}$$

By equation (C.1), this implies

$$\begin{aligned} M_F(y) &> \frac{1}{1-\delta} - \left( \sum_{t=0}^{n-1} \delta^t x_t + \delta^n \left( x_n + \delta M_L(x_n, B) \right) \right) \\ &= \frac{1}{1-\delta} - \left( \sum_{t=0}^{n-1} \delta^t x_t + \delta^n M_L(x_{n-1}, B) \right) \\ &\vdots \\ &= \frac{1}{1-\delta} - M_L(y, B). \end{aligned}$$

But this is a contradiction to the assumption that $y \notin \mathscr{W}^0(B)$. Hence $F$ does not want to preempt. QED.

**Lemma C.3.** *$F$ has no incentive to buy or demand peace when $y \in \mathscr{W}^n(B)$.*

*Proof.* Suppose that $F$ had the incentive to demand peace. That is, suppose there exists $x \in \mathscr{P}(B)$ such that $x < \chi(y,B)$ and $x + V_F(\sigma^*; x, B) > V_F(\sigma^*; y, B)$. $x < \chi(y,B)$ implies that $h(x,B) < h(\chi(y,B),B) = M_L(y,B)$, however, so that no such $x$ can exist.

The remainder of the proof is by induction. Let $y \in \mathscr{W}^1(B)$ and let $x = \chi(y,B)$ so that $x + \delta M_L(x,B) = M_L(y,B)$. Then

$$M_F(y) + M_L(y,B) \leq \frac{\delta}{1-\delta} \quad \text{and} \quad M_F(x) + M_L(x,B) > \frac{1}{1-\delta}.$$

Suppose by way of contradiction that there were $x' > x$, $x' \in \mathscr{P}(B)$, such that

$$1 - x' + \delta V_F(\mathbf{x}(x')) > 1 - x + \delta M_F(x). \tag{C.4}$$

By Lemma C.1, $V_F(\mathbf{x}(x')) = {}^1\!/_{1-\delta} - M_L(x',B)$ and so inequality (C.4) implies

$$1 - x' + \delta \left[ \frac{1}{1-\delta} - M_L(x',B) \right] > 1 - x + \delta M_F(x,B). \tag{C.5}$$

Because $y \in \mathscr{W}^1(B)$, $x \in \mathscr{W}^0(B)$ so that $1/1-\delta - M_L(x,B) < M_F(x)$. Thus inequality (C.5) implies

$$1 - x' + \delta\left[\frac{1}{1-\delta} - M_L(x',B)\right] > 1 - x + \delta\left[\frac{1}{1-\delta} - M_L(x,B)\right]. \tag{C.6}$$

Inequality (C.6) in turn simplifies to

$$x' + \delta M_L(x',B) < x + \delta M_L(x,B) = M_L(y,B).$$

This implies that $L$ rejects an offer of $x'$ today. But this is a contradiction: by Lemma C.2, $F$ prefers to wait a period before going to war.

Suppose now that we have shown that $F$ does not wish to buy peace at all $y \in \mathscr{W}^{n'}(B)$ for $n' < n$. We want to show that $F$ also prefers not to buy peace for all $y \in \mathscr{W}^n(B)$. Consider two cases:

*Case* 1. $F$ is considering a deviation to some $x' \in \mathscr{P}(B)$ such that $x' > y'$ for all $y' \in \mathscr{W}^{n-1}(B)$. This implies that $x'$ was available as a deviation for all $y' \in \mathscr{W}^{n-1}(B)$ and because $F$ does not want to buy peace for any $y' \in \mathscr{W}^{n-1}(B)$, we must have

$$\left[\frac{1}{1-\delta} - M_L(x',B)\right] \leq V_L(\sigma^*; y', B)$$

for all $y' \in \mathscr{W}^{n-1}(B)$. But

$$V_L(\sigma^*; y, B) > \max_{y' \in W^{n-1}(B)} V_L(\sigma^*; y', B).$$

*Case* 2. $F$ is considering a deviation to some $x' \in \mathscr{P}(B)$ such that $x' < y'$ for all $y' \in \mathscr{W}^{n-1}(B)$. In particular, $x' < x_0(y)$, implying that $V_L(x', x_0(x'), x_1(x'), \ldots; y, B) < M_L(y,B)$.[72] Thus $L$ will reject the offer and $F$ prefers not to deviate to $x'$.

[72]This requires that $x_0(y) < 1/2$, which is guaranteed by $y \in W^n(B)$ for $n > 1$.

$\square$

Note that Lemmas C.2 and C.3 require the assumption that $\chi(y,B)$ exists for all $y$, or, if it does not, that the set of all $y$ where it does not exist is a subset of $\mathscr{W}^0(B)$. See Remark C.1 below for a discussion of what happens when this assumption fails.

*Proof of Proposition 2.1.* That $\sigma^*$ is an equilibrium follows from Lemmas C.1, C.2, and C.3. $\square$

*Remark* C.1 (When is $\sigma_F^*$ well-defined?). It is possible that $\sigma_F^*$ is not well-defined. This is the case when there does not exist an $x$ solving $h(x,B) = m_L(y,B)$. Let $\mathscr{N}(B) = \{y | \nexists x, h(x,B) = m_L(y,B)\}$. To this point, we have been assuming that $\mathscr{N}(B) \subset \mathscr{W}^0(B)$. It will be helpful to define the following two conditions:

$$\text{Condition 1: } \mathscr{N}(B) \neq \emptyset$$
$$\text{Condition 2: } \mathscr{N}(B) \not\subseteq \mathscr{W}^0(B)$$

We first comment on the technical aspects of when Condition 1 and Condition 2 hold. We then comment on the viability of $\sigma^*$ as an equilibrium strategy profile under Condition 2.

From a mathematical standpoint, $\mathscr{N}(B) \neq \emptyset$ only if $\arg\max_y m_L(y,B) < y^*(B)$.[73] This occurs only if $B$ is sufficiently greater than 1, which ensures that $m_L(y,B)$ is sufficiently concave. However, for $B \gg 1$, it also turns out that $\mathscr{N}(B) \subset \mathscr{W}^0(B)$. It is straightforward to verify that Lemmas C.1-C.3 hold whether $\mathscr{N}(B) = \emptyset$ or $\emptyset \neq \mathscr{N}(B) \subset \mathscr{W}^0(B)$. Thus $\sigma^*$ is potentially not an equilibrium only if Condition 2 holds.

If Condition 2 holds, it is convenient to redefine the initial war set as $\tilde{\mathscr{W}}^0(B) = \mathscr{W}^0(B) \cup \mathscr{N}(B)$. The following lemma establishes that this is indeed justified:

[73] Recall that $y^*(B)$ is defined $y$ such that $m_L(y,B) = h(y,B)$. It also happens to be the case that $y^*(B) = m_L(y^*(B),B)$.

*Lemma* C.4. *For all $y \in \mathscr{N}(B) \backslash \mathscr{W}^0(B)$, war must occur immediately.*

*Proof.* Suppose instead that there were a sequence of offers that $F$ could make that did not result in war immediately. That is, there is such that $v_L(x_0, x_1, \ldots; \cdot) > m_L(y, B)$. There are two cases:

*Case* 1. $x_n \in \mathscr{N}(B) \backslash \mathscr{W}^0(B)$ for all $n$. This implies that $x_n \leq \sup \mathscr{N}(B) \equiv \bar{x}$ for all $n$. Then $v_L(x_0, x_1, \ldots; \cdot) \leq v_L(\bar{x}, \bar{x}, \ldots; \cdot) \leq y^*(B) < m_L(y, B)$, a contradiction.

*Case* 2. There is $n$ such that $x_{n-1} \in \mathscr{N}(B) \backslash \mathscr{W}^0(B)$ and $x_n \notin \mathscr{N}(B) \backslash \mathscr{W}^0(B)$. Then either (i) $x_n \notin \mathscr{W}^0(B)$, in which case we have that $v_L(x_{n+1}, x_{n+2}, \ldots; \cdot) = m_L(x_n, B)$ and that $v_L(x_n, x_{n+1}, \ldots; \cdot) = x_n + m_L(x_n, B)$; or (ii) $x_n \in \mathscr{W}^0(B)$, $v_L(x_{n-1}, x_n; \cdot) = x_n + m_L(x_n, B)$. In either case, $x_n$ cannot exist by definition.

□

Thus we are justified in treating $\mathscr{N}(B)$ as part of the initial war set.

**Proof of Proposition 2.2**

*Proof.* We treat each case separately:

1. $B = 0$ is always peaceful, establishing existence of peaceful outcomes. To see the existence of $\bar{B}$, note that

$$y^*(B) = \max_{y \in [0,1]} \{y | m_L(y, B) = h(y, B)\} = \max_{y \in [0,1]} \{y | m_L(y, B) = y\}.$$

It is straightforward to show that $0 \in \{y | m_L(y, B) = y\}$ for all $B$ and that there is $y' \neq 0$, $y' \in \{y | m_L(y, B) = y\}$ if and only if

$$\frac{\partial}{\partial y} m_L(y, B)|_{y=0} > 1.$$

Because

$$\frac{\partial}{\partial B}\left(\frac{\partial}{\partial y}m_L(y,B)|_{y=0}\right) > 0 \quad \text{and} \quad \lim_{B\to\infty}\frac{\partial}{\partial y}m_L(y,B)|_{y=0} = \infty$$

$\bar{B}$ exists.

2. That there exist $y$ and $B$ leading to war immediately follows from a stronger statement: namely, that for all $y \in (0,1)$ there is $B$ such that $y \in W^0(B)$. To see this, simply pick $B > B^\dagger(y)$.[74]

3. War in finite time must occur for any $(y,B)$ such that neither case 1 nor case 2 holds.

□

**Proof of Proposition 2.3**

*Proof.* That $B^*(y) > 1$ for all $y$ follows from Lemma 2.1, which says that there is $B^\dagger > 1$ such that

$$M_L(y,B^\dagger) + M_F(y) \leq \frac{1}{1-\delta}$$

for all $y$. This in turn implies that $B^\dagger$ is peaceful. It is straightforward to show that for any two biases $B$ and $B'$ that lead to peaceful outcomes in state $y$, $B > B'$ if and only if $V_H(B,y) > V_H(B',y)$. $V_H(B^\dagger;y)$ is therefore a lower bound on the payoff that $H$ can achieve in equilibrium. Therefore it must be that $B^*(y) > B^\dagger > 1$. □

**Proof of Proposition 2.4**

Suppose the current state is $y$ and consider a strategy profile $\tilde{\sigma}$ defined as follows:

[74] Recall from Lemma 2.1 that $B^\dagger(y)$ is defined to be the minimal bias such that

$$M_L(y,B^\dagger) + M_F(y) = \frac{1}{1-\delta}$$

for a given $y$.

1. A leader of bias $B$ accepts any offer $x$ such that $x \geq M_L(y,B)$.

2. Given $(y,B)$, $F$ offers $x = M_L(y,B)$ if

$$1 - M_L(y,B) + \delta M_F(M_L(y,B)) \geq M_F(y)$$

and $x = 0$ otherwise.

3. $H$ chooses $\tilde{B}(y)$ satisfying

$$(1-\delta C)(1 - M_L(y,B)) = m_F(y)$$

for each $y$.

The following lemma establishes that $B^{**}(y)$ exists and is in fact the unique Markovian rule for choosing a bias that gives $F$ his bias in each period.

**Lemma C.5.** *Consider a mapping $y \mapsto B(y)$ and a sequence $(x_0, x_1, \ldots)$ defined by $x_n = M_L(x_{n-1}, B(x_{n-1}))$. There is a unique mapping $\tilde{B}(y)$ such that this sequence satisfies*

$$V_F(x_n, x_{n+1}, \ldots) = M_F(x_{n-1}) \tag{C.7}$$

*for all $n$.*

*Proof.* Rewrite $V_F(x_n, x_{n+1}, \ldots) = M_F(x_{n-1})$ as

$$1 - x_n + \delta M_F(x_n) = M_F(x_{n-1})$$

or, equivalently,

$$\left(1 + \frac{\delta(1-C)}{1-\delta}\right)(1-x_n) = \frac{m_F(x_{n-1})}{1-\delta} \Leftrightarrow (1-\delta C)(1-x_n) = (1-C)m_F(x_{n-1}).$$

Letting $y = x_{n-1}$ and $x_n = M_L(y,B)$, equation (C.7) is equivalent to:

$$(1-\delta C)(1-M_L(y,B)) = m_F(y). \tag{C.8}$$

Define $\alpha(y,B) = (1-\delta C)(1-M_L(y,B))$. Then $\alpha(y,0) > m_F(y)$ for all $y \in (0,1)$, which, together with the fact that $\alpha(y,B)$ is strictly decreasing in $B$ for all $y \in (0,1)$ implies that there is a unique $\tilde{B}(y) > 0$ satisfying equation (C.8). □

One can verify directly that

$$\tilde{B}(y) = B^{**}(y) \equiv \frac{(1-\delta)C + \delta C(1-C)y + G\left[p(y)(1-y) - \delta C y(1-p(y))\right]}{(1-\delta C)G(1-y)p(y)}$$

and thus that $\tilde{\sigma}$ as defined above is equivalent to $\sigma^{**}$ as defined in Section 2.4.1.

*Proof of Proposition 2.4.* First, note that $L$'s strategy is optimal regardless of what $F$ and $H$ do. Second, note that $F$'s continuation payoff under $\sigma^{**}$ is $M_F(x)$ where $x$ is his offer.[75] Thus his payoff to any offer is decreasing in $x$. This implies that, given $(y,B)$, $F$'s best response to $\sigma^{**}_{-F}$ is to offer $x = M_L(y,B)$, conditional on that offer having payoff at least $M_F(y)$; that is, conditional on

$$1 - M_L(y,B) + \delta M_F(M_L(y,B)) \geq M_F(y).$$

Finally, $H$'s best response to $\sigma^{**}_{-H}$ is to minimize $F$'s payoff, conditional on not going to war. But Lemma C.5 implies that $V_F(\sigma^{**};y) = M_F(y)$ for all $y$; thus $\sigma^{**}_H$ is a best response to $\sigma^{**}_{-H}$. These three points imply that $\sigma^{**}$ is an equilibrium.

That $\sigma^{**}$ is the unique Markov perfect equilibrium also follows from Lemma C.5 and the fact the game is zero sum conditional on war not occurring: $B^{**}$ is the unique rule

[75] This follows from Lemma C.5.

choosing a credible sequence of biases that gives $F$ his minmax payoff in each period and does not lead to war. $\square$

## C.3 Chapter 3 Proofs

**Proof of Lemma 3.1** If the responding potential ally, $j$, rejects, this player receives $v_j(s_0)$. Therefore the proposer optimizes by setting $\phi_i(s_{ij})$ such that

$$u_j(s_{ij},\rho) = \phi_i(s_{ij})u_{ij}(s_{ij}) = v_j(s_0).$$

This is always possible since by A1.2 and $v_i(s_0) \geq 0$, $w_{ij}(s_{ij}) > v_j(s_0)$ which implies that $u_{ij}(s_{ij}) > v_j(s_0)$. Since $\phi_i(s_{ij}) \in [0,1]$ and both $u_{ij}(s_{ij})$ and $v_j(s_0)$ are positive, a solution always exists. Therefore, with probability $p(r_i, r_j)$, $i$ receives $u_{ij}(s_{ij}) - v_j(s_0)$. It must be the case that $u_{ij}(s_{ij}) - v_j(s_0) > v_i(s_0)$ directly by A1.2. When $i$ is the responder, by the same logic, $i$ will receive $v_i(s_0)$. Therefore, the expected value of alliance $ij$ is as stated in Lemma 3.1. QED.

**Proof of Lemma 3.2** By A1.3, $v_i(s_0) > u_i(s_{jz},\rho)$ for all $\rho$. From A1.2, $u_{ij}(s_{ij}) - v_j(s_0) > v_i(s_0)$. By Lemma 3.1, $E_0 u_i(s_{ij},\rho)$ is a linear combination of two values, both of which have been shown to be larger than $u_i(s_{jz},\rho)$ for all $\rho$. QED.

**Proof of Proposition 3.1** First consider the subset $S' \subset S$, where $s_0 \notin S'$, but all other states $s \in S$ are included in $S'$. I first show that there exists a stable state in $S'$.

Without an intra-alliance transfer, $i$ has an ordering over alliances WLOG such that $u_i(s_{ij},\rho) \geq u_i(s_{iz},\rho)$. If $u_j(s_{ij},\rho) \geq u_j(s_{jz},\rho)$ as well, then $s_{ij}$ is stable. So assume that $u_j(s_{jz},\rho) > u_j(s_{ij},\rho)$. If it is the case that $u_z(s_{jz},\rho) \geq u_z(s_{iz},\rho)$, then $s_{jz}$ is stable. So assume that $u_z(s_{iz},\rho) > u_z(s_{jz},\rho)$. Now consider a transfer, $x_i$, so that $u_j(s_{jz},\rho) =$

$u_j(s_{ij},\rho)+x_i$. If then $u_i(s_{ij},\rho) \geq u_i(s_{iz},\rho)+x_i$, then $s_{ij}$ is stable with this transfer. So assume that $u_i(s_{ij},\rho) < u_i(s_{iz},\rho)+x_i$ or $u_i(s_{ij},\rho) < u_i(s_{iz},\rho)+u_j(s_{jz},\rho)-u_j(s_{ij},\rho)$. Furthermore, assume that neither $s_{jz}$ nor $s_{iz}$ can be made stable with an inter-alliance transfer. This leads to the following three inequalities:

$$u_i(s_{ij},\rho) < u_i(s_{iz},\rho)+u_j(s_{jz},\rho)-u_j(s_{ij},\rho)$$
$$u_j(s_{jz},\rho) < u_j(s_{ij},\rho)+u_z(s_{iz},\rho)-u_z(s_{jz},\rho)$$
$$u_z(s_{iz},\rho) < u_z(s_{jz},\rho)+u_i(s_{ij},\rho)-u_i(s_{iz},\rho)$$

Rearranging and plugging the third inequality into the second gives

$$u_j(s_{jz},\rho) < u_j(s_{ij},\rho)+u_z(s_{jz},\rho)+u_i(s_{ij},\rho)-u_i(s_{iz},\rho)-u_z(s_{jz},\rho).$$

Plugging this into the first gives

$$u_i(s_{ij},\rho) < u_i(s_{iz},\rho)+u_j(s_{ij},\rho)+u_z(s_{jz},\rho)+u_i(s_{ij},\rho)-u_i(s_{iz},\rho)-u_z(s_{jz},\rho)-u_j(s_{ij},\rho).$$

Canceling terms gives $u_i(s_{ij},\rho) < u_i(s_{ij},\rho)$ a contradiction. Therefore in all possible cases, there exists a stable state $s$.

Now I show that this extends to the set $S$.

If $v(s_0)$ is such that all players order $s_0$ below the states in which they are in alliances, then there exists a stable state by the above argument. If two players rank $s_0$ highest, then $s_0$ is stable (three players cannot rank $s_0$ highest by A1.2). WLOG, let $z$ rank it second while all others rank it third. The only way this situation is not stable is if

$$u_i(s_{ij},\rho) < u_i(s_{iz},\rho) + u_j(s_{jz},\rho) - u_j(s_{ij},\rho)$$

$$u_j(s_{jz},\rho) < u_j(s_{ij},\rho) + u_z(s_{iz},\rho) - v_z(s_0)$$

$$u_z(s_{iz},\rho) < v_z(s_0) + u_i(s_{ij},\rho) - u_i(s_{iz},\rho)$$

Plugging in as before leads to the same absurdity. WLOG, let $j$ and $z$ rank $s_0$ second, while $i$ ranks it third. This is not stable if

$$u_i(s_{ij},\rho) < u_i(s_{iz},\rho) + u_j(s_{jz},\rho) - v_j(s_0)$$

$$u_j(s_{jz},\rho) < v_j(s_0) + u_z(s_{iz},\rho) - v_z(s_0)$$

$$u_z(s_{iz},\rho) < v_z(s_0) + u_i(s_{ij},\rho) - u_i(s_{iz},\rho)$$

Again, plugging in as before leads to a contradiction. If all three players rank $s_0$ second, then stability holds unless

$$u_i(s_{ij},\rho) < v_i(s_0) + u_j(s_{jz},\rho) - v_j(s_0)$$

$$u_j(s_{jz},\rho) < v_j(s_0) + u_z(s_{iz},\rho) - v_z(s_0)$$

$$u_z(s_{iz},\rho) < v_z(s_0) + u_i(s_{ij},\rho) - v_i(s_0)$$

Rearranging and plugging the third inequality into the second gives $u_j(s_{jz},\rho) < v_j(s_0) + u_z(s_{iz},\rho) - u_z(s_{iz},\rho) + u_i(s_{ij},\rho) - v_i(s_0)$. Rearranging and plugging this into the first inequality gives $u_i(s_{ij},\rho) < v_i(s_0) + u_j(s_{jz},\rho) - u_j(s_{jz},\rho) + u_z(s_{iz},\rho) - u_z(s_{iz},\rho) + u_i(s_{ij},\rho) - v_i(s_0)$. Canceling terms leaves as before the contradiction$u_i(s_{ij},\rho) < u_i(s_{ij},\rho)$.

Now, WLOG, assume that $i$ ranks $s_0$ first in these past three permutations. Then the same method leads to the contradiction $v_i(s_0) < v_i(s_0)$ in all three cases. Therefore, no permutation can be unstable without leading to a contradiction. QED.

Before showing Proposition 3.2 it is convenient to prove the following lemma:

**Lemma C.6.** *Under assumptions A1 and A2, $v_i(s_{ij},\rho) > v_i(s_{iz},\rho)$ for all $\rho$ and all players $i$.*

From A2, $v_i(s_{ij},\rho) > v_i(s_0)$ for all $\rho$. By A1.3, $v_i(s_0) > u_i(s_{jz},\rho)$ for all $\rho$. Therefore, $v_i(s_{ij},\rho) > v_i(s_{iz},\rho)$ for all $\rho$. QED.

**Proof of Proposition 3.2** That $\Gamma$ predicts $s_{ij}$ follows from the following logic. If $u_i(s_{ij},\rho) > u_i(s_{iz},\rho)$ and $u_j(s_{ij},\rho) > u_j(s_{jz},\rho)$, then neither $i$ nor $j$ is willing to make a transfer to each other, but neither $i$ or $j$ have a positive deviation to another alliance. $i$ and $j$ do not have a positive deviation to $s_0$ by assumption A2 and in fact of a mutually positive deviation from $s_0$. By Lemma C.6, neither $i$ or $j$ have a positive deviation to states where they are an outside player and, in fact, have a strict deviation from such a state. Furthermore, $s_{ij}$ is uniquely predicted since $i$ has a strict deviation to $s_{ij}$ in state $s_{iz}$ and $j$ has a strict deviation to $s_{ij}$ in state $s_{jz}$. QED.

**Proof of Proposition 3.3** Under $\mathbb{U}$, $u_a(s_{ab},\rho) > u_a(s_{ac},\rho)$ and $u_b(s_{ab},\rho) > u_b(s_{bc},\rho)$. Under $\mathbb{B}$, $u_b(s_{bc},\rho) > u_b(s_{ab},\rho)$ and $u_c(s_{bc},\rho) > u_c(s_{ac},\rho)$. Therefore, the statements follow from Proposition 3.2. QED.

**Proof of Proposition 3.4** U1, U3, B1 and B2 do not result in alliance switching in period 1, therefore there is no additional impact on the non-rising players' initiation of war from the $F$ case.

In U2, by Proposition 3.1, the resource shift results in an alliance shift from $s_{ab}$ to $s_{ac}$. By Lemma C.6, $v_i(s_{ij},\rho) > v_i(s_{iz},\rho)$ for all $\rho$. This implies that $V_b^P(r_1,s_{ab},\rho) >$

$V_b^P(r_1, s_{ac}, \rho)$. This is the only term that changes for $b$ with an alliance shift, therefore if war is chosen by $b$ without an alliance shift, it is chosen when this type of alliance shift is present. The converse does not hold by the strict inequality. Also by Lemma C.6, $c$ strictly benefits from a peaceful shift in the alliance structure. The effect on $a$ is ambiguous since while $a$ benefits from having a stronger alliance partner next period in $c$, $a$ loses by having a stronger outside player to bargain with since $b$ has resources $r_{b1} \geq r_{c0}$.

WLOG, let $b$ be the rising player. In B3, $V_a^P(r_1, s_{ac}, \rho) > V_a^P(r_1, s_{bc}, \rho)$ by Lemma C.6 and the same implication as in the U2 case. This is the only term that changes for $a$ with an alliance shift, therefore if peace is chosen by $a$ without an alliance shift, it is chosen when this type of alliance shift is present. The converse does not hold by the strict inequality. $c$ instigates war in more cases since it is substituting a stronger player to bargain with in both intra- and inter-alliance bargaining. Since the rising player $b$ is always increasing in its peaceful bargain under $F$, $b$ never initiates war under $F$. However, $b$ may initiate war in $\Gamma(D)$ since $s_{ac}$ will form and $V_b^P(r_1, s_{bc}, \rho) > V_b^P(r_1, s_{ac}, \rho)$ from Lemma C.6 and the same implication as in the U2 case. If the shift in alliance outweighs the increase in resources enough, $b$ may prefer to initiate war. An increase in $b$'s resources such that the resource vector moves from $r_0 = \{41, 40, 20\}$ to $r_1 = \{41, 42, 20\}$ is such an example. Plugging shows that it satisfies the war condition with the following values

$$W_b(r_0, s_{bc}) \approx 828 > u_b(r_0, s_{bc}, \tau) + \delta V_b^P(r_1 s_{ac}, \rho) = 69 + 0.95(671) = 707$$

## C.4 Appendix A Proofs

**Proof of Proposition A.1** $R$ offers $x_t = (1-\delta)M_{ST}$ in all periods $t \geq T$. The most $R$ can offer in periods $t < T$ is $x_t = 1$. Let period $T-1$ be reached. The most that $S$ can get by playing $a$ to $x_{T-1} = 1$ today is $1 + \delta \frac{(1-\theta p_{T-1})(1-d)}{1-\delta}$ compared to the war value of

$\frac{(1-p_{T-1})(1-d)}{1-\delta}$ when $w$ is chose. By the proof of Proposition 1.4, $w$ is optimal at $T-1$ since (1.4) holds. This means that if $T-2$ is reached, $w$ is optimal at $T-2$ since $S$ will get its war value at $T-1$ and again by the proof in Proposition 1.4. This logic extends back through backward induction implying $w$ is optimal for $S$ at $t=0$. QED.

**Proof of Proposition A.2** Since (1.4) holds, there is no delay in war. The best peaceful bargain achievable for $S$ has value $\sum_{t=0}^{T-1}\delta^t+\delta^T\left(M_{ST}+\frac{d}{2(1-\delta)}\right)$. Thus there is no positive deviation from $w$ at $t=0$ to a bargain if

$$M_{S0}>\sum_{t=0}^{T-1}\delta^t+\delta^T\left(M_{ST}+\frac{d}{2(1-\delta)}\right)$$
$$\frac{(1-p_0)(1-d)}{1-\delta}>\frac{1-\delta^T}{1-\delta}+\frac{\delta^T(1-\theta^T p_0)(1-d)}{1-\delta}+\frac{\delta^T d}{2(1-\delta)}$$
$$1-p_0>\frac{1-\delta^T}{1-d}+\delta^T(1-\theta^T p_0)+\frac{\delta^T d}{2(1-d)}$$
$$\theta>\frac{1}{\delta}\left[\left(\frac{1}{p_0}\right)\left(p_0-1+\delta^T+\frac{1}{1-d}+\frac{\delta^T(d-2)}{2(1-d)}\right)\right]^{\frac{1}{T}}$$
$$\theta>\frac{1}{\delta}\left(1+\frac{\delta^T(d-2)+2}{2p_0(1-d)}+\frac{\delta^T-1}{p_0}\right)^{\frac{1}{T}}$$
$$\theta>\frac{1}{\delta}\left(1+\frac{\delta^T d-2\delta^T+2+2\delta^T-2d\delta^T-2+2d}{2p_0(1-d)}\right)^{\frac{1}{T}}$$
$$\theta>\frac{1}{\delta}\left(1+\frac{d}{p_0(1-d)}-\frac{d\delta^T}{2p_0(1-d)}\right)^{\frac{1}{T}}$$

If $\theta\leq\frac{1}{\delta}\left(1+\frac{d}{p_0(1-d)}-\frac{d\delta^T}{2p_0(1-d)}\right)^{\frac{1}{T}}$, then $S$ has a positive deviation from war at $t=0$ to the bargain $x_t=1\ \forall\ t<T$. $R$ optimally accepts this bargain in all periods since (1.4) is satisfied implies that $\delta M_{Rt+1}>M_{Rt}$ for all $t<T$. QED.

## C.5 Appendix B Proofs

**Proof of Lemma B.1** If $i$ does not go to war in $t$, then $i$ must receive a credible bargain with a present value such that $V^i(k_t,b)\geq M^i(k_t)$. The same must hold for $j$. Therefore we have $V^i(k_t,b_t)+V^j(k_t,b_t)>\Xi(k_t)$, a contradiction. QED.

**Proof of Lemma B.2** Case 1: Let $\tilde{V}^{(1)i}(k_t,b^*)=\Xi(k_t)-M^j(k_t)$ then Lemma B.2 follows from Lemma B.1.

Case 2: $\tilde{V}^{(1)i}(k_t,b^*)=u^i(k_t,b_t^*)+\delta\sum_{k_{t+1}\in K_{t+1}}q(k_{t+1}|k_t,b_t^*)\max\{M^i(k_{t+1}),\Xi(k_{t+1})-M^j(k_{t+1})\}$

Lemma B.2 follows if $\forall\, b\in\phi^{ex}$: $V^i(k_t,b)\leq\tilde{V}^{(1)i}(k_t,b^*)$. Then for a particular $k_{t+1}$ we have two sub-cases. Either $\max\{M^i(k_{t+1}),\Xi(k_{t+1})-M^j(k_{t+1})\}=M^i(k_{t+1})$ in which case by Lemma B.1, war will occur in state $k_{t+1}$ under rationality. Or, we have

$$\max\{M^i(k_{t+1}),\Xi(k_{t+1})-M^j(k_{t+1})\}=\Xi(k_{t+1})-M^j(k_{t+1})$$

in which case, by definition $V^i(k_{t+1},b)\leq\Xi(k_{t+1})-M^j(k_{t+1})$. Therefore, for any particular $k_{t+1}$ and $b$, the continuation value under $\tilde{V}^{(1)i}(k_t,b^*)$ is higher than that under $V^i(k_t,b)$. The first term, $u^i(k_t,b_t)$, is the same for $\tilde{V}^{(1)i}(k_t,b^*)$ and $V^i(k_t,b)$. Putting these three elements together, for any particular bargain $b$ we have

$$u^i(k_t,b_t)+\delta\sum_{k_{t+1}\in K_{t+1}}q(k_{t+1}|k_t,b_t)\max\{M^i(k_{t+1}),\Xi(k_{t+1})-M^j(k_{t+1})\}\geq V^i(k_t,b)$$

By definition, $b^*$ maximizes the LHS, therefore the LHS under $b^*$ is greater than any bargain in the RHS. QED.

**Proof of Lemma B.3** If the RHS equals $\tilde{V}^{(1)i}(k_t,b^*)$, then war follows from Lemma B.2. Otherwise, Lemma B.3 follows if $\forall\, b\in\phi^{ex}$

$$u^i(k_0,b_0^*)+\delta[\mu_1^{(1)i}(k_0,b^*)+\hat{M}_1^{(1)i}(k_0,b^*)+\hat{M}_1^{(1)ij}(k_0,b^*)]+\delta^2\hat{\Upsilon}_2^i(k_0)\geq V^i(k_0,b)$$

By definition

$$V^i(k_0,b_0) = u^i(k_0,b_0) + \delta[\sum_{k_1 \in K_1^W(b)} q(k_1|k_0,b_0)M^i(k_1) + \sum_{k_1 \in K_1^P(b)} q(k_1|k_0,b_0)V^i(k_1,b_1)]$$

Plugging in for $V^i(k_1,b_1)$, this becomes

$$V^i(k_0,b_0) = u^i(k_0,b_0) + \delta[\sum_{k_1 \in K_1^W(b)} q(k_1|k_0,b_0)M^i(k_1) +$$
$$+\sum_{k_1 \in K_1^P(b)} q(k_1|k_0,b_0)(u^i(k_1,b_1) + \delta[\sum_{k_2 \in K_2^W(b)} q(k_2|k_1,b_1)M^i(k_2)$$
$$+\sum_{k_2 \in K_2^P(b)} q(k_2|k_1,b_1)V^i(k_2,b_2)])]$$

Consider a particular sequence of states $k_0$, $k_1$ and $k_2$ as well as bargains $b_0$, $b_1$ and $b_2$. I then want to show that

$$u^i(k_0,b_0) + \delta[\mu_1^{(1)i}(k_0,b) + \hat{M}_1^{(1)i}(k_0,b) + \hat{M}_1^{(1)ij}(k_0,b)] + \delta^2\hat{\Upsilon}_2^i(k_0) \geq$$
$$u^i(k_0,b_0) + \delta[\sum_{k_1 \in K_1^W(b)} q(k_1|k_0,b_0)M^i(k_1) + \sum_{k_1 \in K_1^P(b)} q(k_1|k_0,b_0)(u^i(k_1,b_1) +$$
$$\delta[\sum_{k_2 \in K_2^W(b)} q(k_2|k_1,b_1)M^i(k_2) + \sum_{k_2 \in K_2^P(b)} q(k_2|k_1,b_1)V^i(k_2,b_2)])]$$

Since the first terms are always the same, we can cancel them immediately. Next, we can see that there three possible cases.

In the first case, we have that $V^i(k_1,b_1) \geq M^i(k_1)$, which implies, by Lemma B.2, that $\Upsilon^i(k_1) \geq M^i(k_1)$ so that neither the LHS or RHS predicts war in period 1. This means that the LHS assigns this case a value of $\delta u^i(k_1,b_1) + \delta^2\Upsilon^i(k_2)$ and the RHS assigns a value of $\delta u^i(k_1,b_1) + \delta^2 \max\{M^i(k_2), V^i(k_2,b_2)\}$. It follows from Lemma B.2 and the definition of $\Upsilon^i(k_2)$ that $\Upsilon^i(k_2) \geq \max\{M^i(k_2), V^i(k_2,b_2)\}$. In case 2, war occurs in period 1 under both $SC(1)$ and rationality in which case the LHS and RHS are equal. In the third case, rationality predicts war at $t = 1$, but $SC(1)$ fails to. For a given bargain $b$, the RHS assigns a value of $\delta M^i(k_1)$ regardless of the realized values of $k_2$ and $b_2$. Meanwhile, the LHS, given $k_1$ and $b$ has an expected value of $\delta u^i(k_1,b_1)+\delta^2\hat{\Upsilon}^i_2(k_0)$. Canceling the $\delta$s, and by the Law of Iterated Expectations, the LHS has exactly the value of $\tilde{V}^{(1)i}(k_1,b^*)$. But, since $i$ did not choose war under $SC(1)$ in this case, so it must be that $\delta u^i(k_1,b_1)+\delta^2\hat{\Upsilon}^i_2(k_0,b)\geq \delta M^i(k_1)$.

Since war occurs in possibly more states under rationality, it may be that the sequence of states do not occur with the same probability in the LHS and the RHS. However, since peace is assumed in period 0, a particular state $k_1$ occurs with the same frequency. Since the third case does not depend on the value of $k_2$ and $b_2$, the differing probabilities of their outcome does not matter. In cases one and two, $SC(1)$ and rationality give the same outcomes at $t=1$, so $k_2$ will occur with the same probability. Since the LHS does not depend on the value of $b_2$, the frequency of bargaining outcomes on the RHS are irrelevant for cases one and two. Therefore, adding all three cases by their frequency implies that the LHS$\geq$RHS for a given bargain $b$. By definition, $b^*$ maximizes the LHS, therefore the LHS under $b^*$ is greater than any bargain in the RHS. QED.

**Proof of Proposition B.1** We are done if we can show that $\tilde{V}^{(n)i}(k_0,b_0^*)\geq V^i(k_0,b)\ \forall\ b\in\phi^{ex}$. This means showing that

$$u^i(k_0,b_0^*)+\sum_{t=1}^{n-1}\delta^t[\mu_t^{(n-1)i}(k_0)+\hat{M}_t^{(n-1)i}(k_0,b_0^*)+\hat{M}_t^{(n-1)ij}(k_0,b_0^*)]+\delta^n\hat{\Upsilon}^i_n(k_0)\geq V^i(k_t,b_t)$$

$\forall\ b\in\phi^{ex}$. I have shown the basis case in Lemma B.3, so assume that $\tilde{V}^{(n-1)i}(k_t,b_t)\geq V^i(k_t,b_t)$ and $\tilde{V}^{(n-1)j}(k_t,b_t)\geq V^j(k_t,b_t)$ or $SC(n-1)$ is a true sufficient condition. If the RHS equals $\tilde{V}^{(n-1)i}(k_t,b_t)$, then war follows from the assumption that the basis case is a sufficient condition.

For a given sequence $b^n$ and $k^n$, under rationality there are two cases. Either war occurs at some time $t\leq n$ or it does not. If war does not occur under rationality then since $SC(n-$

1) is a sufficient condition, war does not occur under $SC(n)$ for any period other than 0 since $SC(n-1)$ is the condition for in every period $0 < t \leq n$ in forming $SC(n)$. Therefore for a sequence in which war does not occur under rationality we have a value of

$$\sum_{t=0}^{n-1} \delta^t u^i(k_t b_t) + \delta^n V^i(k_n, b_n)$$

while the same sequence while deciding on war based on $SC(n)$ is

$$\sum_{t=0}^{n-1} \delta^t u^i(k_t b_t) + \delta^n \Upsilon^i(k_n)$$

The first terms are the same but $\Upsilon^i(k_n) \geq V^i(k_n, b_n)$ by Lemma B.2.

In the case that a sequence $(k^n, b^n)$ leads to war under rationality for $0 < \tau \leq n$ there are two cases under $SC(n)$. $\tau > 0$ by definition of $V^i$. In the first case $SC(n)$ leads to war in period $\tau$ as well in which case the sequences give equal value. In the second case, war occurs under rationality which gives a value of

$$\sum_{t=0}^{\tau-1} \delta^t u^i(k_t b_t) + \delta^\tau M^i(k_\tau)$$

but does occur at $\tau$ not under $SC(n)$ which gives a value of

$$\sum_{t=0}^{\tau-1} \delta^t u^i(k_t b_t) + \delta^\tau \tilde{V}^{(n-1)i}(k_\tau, b_\tau)$$

Since rationality leads to war at $\tau$ but not for $SC(n-1)$ we must have $\tilde{V}^{(n-1)i}(k_\tau, b_\tau) \geq M^i(k_\tau) > V^i(k_\tau, b_\tau)$. Which means that for the truncated sequence $(k^\tau, b^\tau)$ we have $\tilde{V}^{(n)i}(k^\tau, b^\tau) \geq V^i(k^\tau, b^\tau)$.

This means that every sequence $(k^n, b^n)$ gives a value that is either directly $\tilde{V}^{(n)i}(k^n, b^n) \geq V^i(k^n, b^n)$ under no war or the first case of war. Or, under the second war case, since $\tau > 0$, the expectation of every sequence beginning with $(k_0, b_0)$ is such that $\tilde{V}^{(n)i}(k_0, b_0) \geq$

$V^i(k_0, b_0)$.

Since it is the case then that $\tilde{V}^{(n)i}(k_0, b) \geq V^i(k_0, b)\ \forall\ b \in \phi^{ex}$ and $b^*$ maximizes the LHS, it must be the case that $\tilde{V}^{(n)i}(k_0, b_0^*) \geq V^i(k_0, b)\ \forall\ b \in \phi^{ex}$. This shows the induction step. QED.

**Proof of Proposition B.2** I prove each statement in turn.

1. ($SC(n,1)$ sufficient condition for war) If $\phi^{en}(k_t, n) = \emptyset$, then war occurs at $k_t$.

Consider a potentially peaceful bargain $b$. If $\phi^{en}(k_t, n) = \emptyset$, then either $M^i(k_t) > \tilde{V}^{(n)i}(k_t, b)$ or $M^j(k_t) > \tilde{V}^{(n)j}(k_t, b)$ or both. Thus, by Proposition B.1, war occurs under bargain $b$ at $k_t$. Since $b$ was chosen arbitrarily, this holds for all bargains $b \in \phi^{ex}$.

2. If $\phi^{en}(k_t, n) \neq \emptyset$, then for any peaceful bargain $b$ at $k_t$ it is the case that $b \in \phi^{en}(k_t, n)$.

Pick a $b \notin \phi^{en}(k_t, n)$, so then $M^i(k_t) > \tilde{V}^{(n)i}(k_t, b)$ or $M^j(k_t) > \tilde{V}^{(n)j}(k_t, b)$ or both. Which implies by Proposition B.1 that $M^i(k_t) > V^i(k_t, b)$ or $M^j(k_t) > V^j(k_t, b)$ or both. Thus war occurs at $k_t$. Therefore $b$ is not a peaceful bargain.

3. If war occurs under $SC(n)$, then it occurs under $SC(n,1)$. The converse does not hold.

War occurs under $SC(n)$ when there exists a player $i$ such that $M^i(k_t) > \tilde{V}^{(n)i}(k_t, b)$. When this occurs, by definition of $\phi^{en}(k_t, n)$ we have $\phi^{en}(k_t, n) = \emptyset$. By statement 1, war occurs under $SC(n,1)$. The converse does not hold by the following counterfactual:

Assume players labeled $H$ and $F$ where $i$ and $j$ are generic players. Let $b_t^i \in [0,1]$ and $b_t^H + b_t^F \leq 1$. Let $u^i(b_t^i) = b_t^i$. Assume a costly war process so that there exists a bargain that Pareto dominates the war outcome. Let $\delta = 0.95$ so that $\Xi_t = 20$. Let $M^H(k_t) = 8$ and $M^F(k_t) = 8$ for all $t$. Imagine that the issue bargained over is completely indivisible so that we have a set $\phi^{ex}$ such that $b_t^H = 1$ and $b_t^F = 0$ for all $t$ or $b_t^H = 0$ and $b_t^F = 1$ for all $t$. No other bargains are possible. From either $H$ or $F$'s perspective $b_t^* = 1$ for all $t$, thus we have $\tilde{V}^{(n)i}(k_t, b^*) = 20$ which is greater than the minmax value of 8. Therefore neither chooses war under $SC(n)$. Now consider $SC(n,1)$. For both possible bargains there exists one player with $\tilde{V}^{(n)i}(k_t,b)=0$, which means that at least one player prefers war under both bargains so that $\phi^{en}(k_t,n)=\emptyset$. Thus war occurs under $SC(n,1)$. QED.

**Proof of Lemma B.4** First we show that $\Lambda^{(n,2)i}(k_0,b)\geq V^i(k_0,b)$. This means showing that

$$v_0^{(n,1)i}(k_0,b)+\delta[P_1^{(n,1)i}(k_0,b)+W_1^{(n,1)i}(k_0,b)]\geq$$
$$u^i(k_0,b_0)+\delta[\textstyle\sum_{k_1\in K_1^W(b)}q(k_1|k_0,b_0)M^i(k_1)+\sum_{k_1\in K_1^P(b)}q(k_1|k_0,b_0)V^i(k_1,b)]$$

By definition $v_0^{(n,1)i}(k_0,b)=u^i(k_0,b_0)$. For a given $k_1$ we have by definition $W_1^{(n,1)i}(k_0,b)=M^i(k_1)$. By definition of $P_1^{(n,1)i}(k_0,b)$ and Proposition B.1 we have that for a given $k_1$, $P_1^{(n,1)i}(k_0,b)\geq V^i(k_1,b)$. Of course there maybe more $k_1$ that result in war on the RHS then the LHS. However, in each of these cases the LHS assigns a value of $\tilde{V}^{(n)i}(k_1,b)\geq M^i(k_1)$ otherwise war would have been chosen at $k_1$.

Now I show each statement in turn.

1. ($SC(n,2)$ sufficient condition for war) If $\phi^{en}(k_0,n,2)=\emptyset$, then war occurs at $k_t$.

Consider a potentially peaceful bargain $b$. If $\phi^{en}(k_0,n)=\emptyset$, then either $M^i(k_0)>\Lambda^{(n,2)i}(k_0,b)$ or $M^j(k_0)>\Lambda^{(n,2)i}(k_0,b)$ or both. Since $\Lambda^{(n,2)i}(k_0,b)\geq V^i(k_0,b)$ this means that war occurs at $k_0$ under bargain $b$. Since $b$ was chosen arbitrarily, this holds for all bargains $b\in\phi^{ex}$.

2. If $\phi^{en}(k_0,n,2)\neq\emptyset$, then for any peaceful bargain $b$ at $k_0$ it is the case that $b\in\phi^{en}(k_0,n,2)$.

Pick a $b\notin\phi^{en}(k_0,n,2)$, so then $M^i(k_0)>\Lambda^{(n,2)i}(k_0,b)$ or $M^j(k_0)>\Lambda^{(n,2)i}(k_0,b)$ or both. This means that war occurs at $k_0$ and $b$ is not a peaceful bargain.

3. If war occurs under $SC(n,1)$, then it occurs under $SC(n,2)$. The converse does not hold.

The first part follows if $\tilde{V}^{(n)i}(k_0,b) \geq \Lambda^{(n,2)i}(k_0,b)$ and therefore $\Lambda^{(n,2)i}(k_0,b)$ provides a tighter bound. $\Lambda^{(n,2)i}(k_0,b)$ uses the condition $SC(n,1)$ to determine which states $k_1$ result in war whereas $\tilde{V}^{(n)i}(k_0,b)$ uses condition $SC(n-1)$. We know from Proposition B.2 that $SC(n,1)$ is a broader condition than $SC(n-1)$. This means that in every case $k_1$ that the LHS assigns a war value $M^i(k_1)$, the RHS also assigns the same value. Furthermore, there are case where the RHS assigns a war value but the LHS does not. In these cases the expected value of the RHS is less than or equal to the LHS otherwise $i$ would have chosen war. In all other cases the RHS assigns a value of $\tilde{V}^{(n)i}(k_1,b)$. These are all peaceful states for the LHS as well, so we know that the expected value of the LHS must be $\geq \tilde{V}^{(n-1)i}(k_1,b)$ otherwise war would have been chosen under $SC(n-1)$. Since by definition $\tilde{V}^{(n)i}(k_1,b) \leq \tilde{V}^{(n-1)i}(k_1,b)$, these cases also assign a value to the RHS that is less than or equal to the LHS. Since all cases are satisfied and $q(k_1|k_0)$ is the same for both sides in the absence of war at $k_0$ we have $\tilde{V}^{(n)i}(k_0,b) \geq \Lambda^{(n,2)i}(k_0,b)$.

The second part follows from the following counterfactual with the same setup as in Proposition B.2 with the following adjustments.Let $k_0$ be such that $M^H(k_0) = 9.5$ and $M^F(k_0) = 9.5$. If war does not occur, then let $q(k_1^{'}|k_0) = 0.5$ and $q(k_1^{''}|k_0) = 0.5$. For $k_1^{'}$ and $k_1^{''}$, assume that $M^H(k_1^{'}) = 8$, $M^H(k_1^{''}) = 8$, $M^F(k_1^{'}) = 8$, and $M^F(k_1^{''}) = 8$ and remain at 8 for all periods $t \geq 1$. Assume that at $k_0$ any $b_0$ such that $b_0^H + b_0^F \leq 1$ is part of $\phi^{ex}(k_0)$ and at $k_1^{'}$ any $b_t^H + b_t^F \leq 1$ for all $t \geq 1$ is part of $\phi^{ex}(k_1^{'})$. However, for $\phi^{ex}(k_1^{''})$ assume that bargains are restricted as in the counterfactual in Proposition B.2.

Consider a bargain $\hat{b}$ such that $b_0^H = 0.5$ then if $k_1^{'}$ is realized, $\hat{b}_t^H = 0$ in for all $t \geq 1$ while if $k_1^{''}$ is realized, $\hat{b}_t^H = 1$ in for all $t \geq 1$. This means that $\tilde{V}^{(n)H}(k_0,\hat{b}) = 10$ which is greater than $H$'s minmax value in 0. Similarly $F$ also has $\tilde{V}^{(n)F}(k_0,\hat{b}) = 10$ given a symmetric bargain. Therefore it is the case that $\hat{b} \in \phi^{en}(k_t,n)$. This means that $SC(n,1)$

does not predict war in this situation.

However, $SC(n,2)$ does predict war here. It is immediate that $\hat{b} \notin \phi^{en}(k_0,n,2)$ since from the counterfactual in Proposition B.2 it is known that $\hat{b} \notin \phi^{en}(k_1^{''},n)$. Also from the previous counterfactual, it is known that war must occur at $k_1^{''}$. This means that both $H$ and $F$ assign an undiscounted value of 8 to state $k_1^{''}$. In state $k_1^{'}$ we need only check the case where $H$ and $F$ split the bargain equally giving them an undiscounted value of $P_1^{(n,1)i}(k_0,b) = 10$ for both players. One need only check this case since this maximizes the minimum bargain value for the players. Since their minmax value at $k_0$ and $k_1^{''}$ are equal this is the least likely bargain for $k_1^{'}$ to cause war. For the same reason we can assume an equally split bargain at $k_0$ giving a value of $v_0^{(n,1)i}(k_0,b) = 0.5$ for both players. This means that for player $H$ we have value $\Lambda^{(n,2)H}(k_0,b) = 0.5+0.95[(0.5)(10)+(0.5)(8)] = 9.05$, but since $M^H(k_0) = 9.5 > 9.05$ this is not part of $\phi^{en}(k_0,n,2)$. Since this was the most likely bargain to be part of this set we can conclude that $\phi^{en}(k_0,n,2) = \emptyset$ and that war condition $SC(n,2)$ is satisfied. QED.

**Proof of Proposition B.3** First we show that $\Lambda^{(n,m)i}(k_0,b) \geq V^i(k_0,b)$. We know that by definition

$$V^i(k_0,b) = u^i(k_0,b_0) + \delta[\sum_{k_1 \in K_1^W(b)} q(k_1|k_0,b_0)M^i(k_1) + \sum_{k_1 \in K_1^P(b)} q(k_1|k_0,b_0)V^i(k_1,b)]$$

Now imagine repeatedly plugging in for $V^i(k_t,b)$ in the last term until the terminal term is $V^i(k_m,b)$. Imagine the hypothetically perfect sufficient condition that is achieved under rationality. Call this condition "$r$" so that $v_t^{(n,r)i}(k_0,b)$, $W_t^{(n,r)i}(k_0,b)$, and $P_t^{(n,r)i}(k_0,b)$ denote the expected peaceful bargains instantaneous values, war values, and continuation values respectively under rationality for a given bargain. Using this notation and plugging in for $\Lambda^{(n,m)i}(k_0,b)$, $\Lambda^{(n,m)i}(k_0,b) \geq V^i(k_0,b)$ becomes

$$\sum_{t=1}^{m-1} \delta^t [v_{t-1}^{(n,m-1)i}(k_0,b) + W_t^{(n,m-1)i}(k_0,b)] + \delta^{m-1} P_{m-1}^{(n,m-1)i}(k_0,b) \geq$$
$$u^i(k_0,b_0) + \sum_{t=1}^{m-2} \delta^t [v_t^{(n,r)i}(k_0,b) + W_t^{(n,r)i}(k_0,b)] + \delta^{m-1} P_{m-1}^{(n,r)i}(k_0,b)$$

From inspection of the left and right hand sides of this inequality, we can determine that the only difference is that the condition $r$ is broader than the condition $SC(n,m-1)$. Therefore, if for a given sequence $k^T$, war occurs at time $T$ on the LHS, it also occurs on the RHS. If war never occurs on the RHS for a given sequence $k^{m-1}$, then it also does not occur on the LHS since $SC(n,m-1)$ is a sufficient condition. There are however sequence where $k^T$ where war occurs at state time $T$ on the RHS, but at some later date or never on the LHS. This means that the RHS assigns a value to sequence $k^T$ of

$$\sum_{t=0}^{T-1} \delta^t u^i(k_t b_t) + \delta^T M^i(k_T)$$

For $\Lambda^{(n,m)i}(k_0,b)$ (the LHS), the first $T-1$ periods also give a value of $\sum_{t=0}^{T-1} \delta^t u^i(k_t b_t)$, but in period $T$ a war value is not assigned. Since under $\Lambda^{(n,m)i}(k_0,b)$, $i$ could have chosen war at $T$, it must be that the expectations of $\Lambda^{(n,m)i}(k_0,b)$ at $T$ is greater than $\delta^T M^i(k_T)$. By the Law of Iterated Expectations, it must be the case that for all of these types of sequences $\Lambda^{(n,m)i}(k_0,b)$ assigns a value that is greater than or equal to $V^i(k_0,b)$ (the RHS). Since in all other sequences they are equivalent we have $\Lambda^{(n,m)i}(k_0,b) \geq V^i(k_0,b)$. This now allows us to use $\Lambda^{(n,m)i}(k_0,b)$ as a bound on $V^i(k_0,b)$ in forming the $SC(n,m)$ condition for war.

1. ($SC(n,m)$ sufficient condition for war) If $\phi^{en}(k_0,n,m) = \emptyset$, then war occurs at $k_t$.

Consider a potentially peaceful bargain $b$. If $\phi^{en}(k_0,n,m) = \emptyset$, then either $M^i(k_0) > \Lambda^{(n,m)i}(k_0,b)$ or $M^j(k_0) > \Lambda^{(n,m)i}(k_0,b)$ or both. Since $\Lambda^{(n,m)i}(k_0,b) \geq V^i(k_0,b)$ this means that war occurs at $k_0$ under bargain $b$. Since $b$ was chosen arbitrarily, this holds for all bargains $b \in \phi^{ex}$.

2. If $\phi^{en}(k_t,n,m) \neq \emptyset$, then for any peaceful bargain $b$ at $k_t$ it is the case that $b \in \phi^{en}(k_t,n,m)$.

Pick a $b \notin \phi^{en}(k_t,n,m)$, so then $M^i(k_0) > \Lambda^{(n,m)i}(k_0,b)$ or $M^j(k_0) > \Lambda^{(n,m)i}(k_0,b)$ or both. This means that war occurs at $k_0$ and $b$ is not a peaceful bargain.

3. If war occurs under $SC(n,m-1)$, then it occurs under $SC(n,m)$.

The prove is finished if I can show that $\Lambda^{(n,m)i}(k_0,b) \geq \Lambda^{(n,m+1)i}(k_0,b)$. Consider $\Lambda^{(n,m+1)i}(k_0,b)$, but where condition $SC(n,m-1)$ instead of $SC(n,m)$. Call this $\tilde{\Lambda}^{(n,m+1)i}(k_0,b)$. Then

$$\tilde{\Lambda}^{(n,m+1)i}(k_0,b) = \sum_{t=1}^{m} \delta^t [v_{t-1}^{(n,m-1)i}(k_0,b) + W_t^{(n,m-1)i}(k_0,b)] + \delta^m P_m^{(n,m-1)i}(k_0,b)$$

Now subtract from this $\Lambda^{(n,m)i}(k_0,b)$ and divide by $\delta^{m-1}$ so that

$$v_{m-1}^{(n,m-1)i}(k_0,b) + \delta W_m^{(n,m-1)i}(k_0,b) + \delta P_m^{(n,m-1)i}(k_0,b) - P_{m-1}^{(n,m-1)i}(k_0,b)$$

For a given $k_{m-1}$, $P_{m-1}^{(n,m-1)i}(k_0,b) = \tilde{V}^{(n)i}(k_{m-1},b)$. Therefore, if

$$\tilde{V}^{(n)i}(k_{m-1},b) \geq v_{m-1}^{(n,m-1)i}(k_0,b) + W_m^{(n,m-1)i}(k_0,b) + \delta P_m^{(n,m-1)i}(k_0,b)$$

I will have shown that $\Lambda^{(n,m)i}(k_0,b) \geq \tilde{\Lambda}^{(n,m+1)i}(k_0,b)$. But this must be true by the same argument as in statement 3 in Lemma B.4 with the only change being the broader condition being used is $SC(n,m-1)$ instead of $SC(n,1)$ in Lemma B.4.

Now in order to complete the proof, I must show that $\tilde{\Lambda}^{(n,m+1)i}(k_0,b) \geq \Lambda^{(n,m+1)i}(k_0,b)$. Since the only difference between these expressions is that the RHS uses a broader war condition, some sequences will occur where the RHS assigns a war value when either, (1) the LHS does not or (2) the LHS assigns war at a later date. When for a particular sequence of states war occurs at $T$ under $\Lambda^{(n,m+1)i}(k_0,b)$ we have

$$\sum_{t=0}^{T-1}\delta^t u^i(k_t b_t)+\delta^T M^i(k_T)$$

For $\tilde{\Lambda}^{(n,m+1)i}(k_0,b)$, the first $T-1$ periods also give a value of $\sum_{t=0}^{T-1}\delta^t u^i(k_t b_t)$, but in period $T$ a war value is not assigned. Since under $\tilde{\Lambda}^{(n,m+1)i}(k_0,b)$, $i$ could have chosen war at $T$, it must be that the expectations of $\tilde{\Lambda}^{(n,m+1)i}(k_0,b)$ at $T$ is greater than $\delta^T M^i(k_T)$. By the Law of Iterated Expectations, it must be the case that for all of these types of sequences $\tilde{\Lambda}^{(n,m+1)i}(k_0,b)$ assigns a value that is greater than or equal to $\Lambda^{(n,m+1)i}(k_0,b)$. Since in all other sequences they are equivalent we have $\tilde{\Lambda}^{(n,m+1)i}(k_0,b)\geq\Lambda^{(n,m+1)i}(k_0,b)$. Which implies by the earlier step that $\Lambda^{(n,m)i}(k_0,b)\geq\Lambda^{(n,m+1)i}(k_0,b)$. QED.

# References

[1] Abreu, Dilip and Faruk Gul. 2000. "Bargaining and Reputation," *Econometrica* 68(1): 85-117.

[2] Abreu, Dilip and David Pearce. 2007. “Bargaining, Reputation, and Equilibrium Selection in Repeated Games with Contracts.” *Econometrica* 75(3): 653-710.

[3] Abreu, Dilip and David Pearce. 2011. “Implementing the Nash Program in Stochastic Games.” Typescript.

[4] Bloch, Francis. 2012. “Endogenous Formation of Alliances in Conflicts.” In *The Oxford Handbook of the Economics of Peace and Conflict* eds. Michelle Garfinkel and Stergious Skaperdas. New York: Oxford University Press, 473-502.

[5] Busch, Lutz-Alexander and Quan Wen. 1995. “Perfect Equilibria in a Negotiation Model.” *Econometrica* 63 (3): 545-565.

[6] Chadefaux, Thomas. 2011. “Bargaining over power: When do shifts in power lead to war?” *International Theory* 3 (2): 228-253.

[7] Chapman, Terrence, Patrick J. McDonald, and Scott Moser. 2012. “The Sword and the Coffers: The Fiscal Foundations of Sustainable International Peace.” Typescript.

[8] Churchill, Winston S. 1949. *The Second World War: Their Finest Hour*. Boston: Houghton Mifflin Company.

[9] Compte, Olivier and Philippe Jehiel. 2002. “On the Role of Outside Options in Bargaining with Obstinate Parties.” *Econometrica* 70(4): 1477-1517.

[10] Copeland, Dale C. 2000. *The Origins of Major War*. Ithaca: Cornell University Press.

[11] Debs, Alexandre and Nuno P. Monteiro. “Known Unknowns: Power Shifts, Uncertainty, and War.” *International Organization*. Forthcoming.

[12] Deudney, Daniel H. 2007. *Bounding Power: Republican Security Theory from the Polis to the Global Village*. Princeton: Princeton University Press.

[13] Dutta, Prajit and Aldo Rustichini. 1993. “A Theory of Stopping Time Games with Applications to Product Innovation and Asset Sales.” *Economic Theory* 3 (4): 743-763.

[14] Fearon, James D. 1995. “Rationalist Explanations for War.” *International Organization* 49 (3): 379-414.

[15] Fearon, James D. 1996. “Bargaining over Objects that Influence Future Bargaining Power.” Typescript.

[16] Fernendez, Raquel and Jacob Glazer. 1991. “Striking for a Bargain Between Two Completely Informed Agents.” *The American Economic Review* 81(1): 240-252.

[17] Fudenberg, Drew and Jean Tirole. 1985. “Preemption and Rent Equalization in the Adoption of New Technology.” *Review of Economic Studies* 52 (3): 383-401.

[18] Garfinkel, Michelle. 2004. “Stable Alliance formation in Distributional Conflict.” *European Journal of Political Economy* 20: 829-852.

[19] Haller, Hans and Steinar Holden. 1990. “A Letter to the Editor on Wage Bargaining.” *Journal of Economic Theory* 52: 232-256.

[20] Jackson, Mathew O. and Massimo Morelli. 2007. “War, Transfers, and Political Bias.” *The American Economic Review*. 97(4): 1353-1373.

[21] Jackson, Mathew O. and Massimo Morelli. 2009. "Strategic Militarization, Deterrence and Wars." *Quarterly Journal of Political Science* 4 (3): 279-313.

[22] Lee, Jihong and Hamid Sabourian. 2007. "Coase Theorem, Complexity, and Transaction Costs." *Journal of Economic Theory* 135(1): 214-235.

[23] Leventoglu, Bahar and Branislav Slantchev. 2007. "The Armed Peace: A Punctuated Equilibrium Theory of War." *American Journal of Political Science* 51 (4): 755-771.

[24] McDonald, Patrick J. 2011. "Complicating Commitment: Free Resources, Power Shifts, and the Fiscal Politics of Preventive War." *International Studies Quarterly* 55: 1095-1120.

[25] Mombauer, Annika. 2001. *Helmuth von Moltke and the Origins of the First World War*. Cambridge: Cambridge University Press.

[26] Muthoo, Abhinay. 1995a. "Bargaining in a Long-term Relationship with Endogenous Termination." *Journal of Economic Theory* 66 (2): 590-598.

[27] Muthoo, Abhinay. 1995b. "On the Strategic Role of Outside Options in Bilateral Bargaining." *Operations Research* 43(2): 292-297.

[28] National Security Council. August 3, 1990. *Meeting of the NSC Meeting*. Meeting notes. Retrieved May 14, 2013 from Margaret Thatcher Foundation Website: www.margaretthatcher.org

[29] Nash, John. 1953. "Two-Person Cooperative Games." *Econometrica* 21(1): 128-140.

[30] Niou, Emerson and Peter Ordeshook. 1986. "A Theory of the Balance of Power in International Systems." *The Journal of Conflict Resolution* 30 (4): 685-715.

[31] Noh, Suk Jae. 2002. "Resource Distribution and Stable Alliances with Endogenous Sharing Rules." *European Journal of Political Economy* 18: 129-151.

[32] Organski, A.F.K. 1968. *World Politics*. New York: Alfred A. Knopf.

[33] Organski, A.F.K and Jacek Kugler. 1980. *The War Ledger*. Chicago: University of Chicago Press.

[34] Paine, S.C.M. 2003. *When Japan Became A World Power: The Sino-Japanese War of 1894-1895*. New York: Cambridge University Press.

[35] Pawlina, Grzegorz and Peter M. Kort. 2006. "Real Options in Asymmetric Duopoly: Who Benefits from Your Comparative Disadvantage?" *Journal of Economics & Management Strategy* 15 (1): 1-35.

[36] Powell, Robert. 1999. *In the Shadow of Power*. Princeton: Princeton University Press.

[37] Powell, Robert. 2004. "The Inefficient Use of Power: Costly Conflict with Complete Information." *American Political Science Review* 98 (2): 231-241.

[38] Powell, Robert. 2006. "War as a Commitment Problem." *International Organization* 60 (1): 169-203.

[39] Powell, Robert. 2012. "Persistent Fighting to Forestall Adverse Shifts in the Distribution of Power." *American Journal of Political Science* 56 (3): 620-637.

[40] Ray, Debraj and Rajiv Vohra. 2013. "Coalition Formation." Typescript.

[41] Rubinstein, Ariel. 1982. "Perfect Equilibria in a Bargaining Model." *Econometrica* 50(1): 97-110.

[42] Schelling, Thomas C. 1963. *The Strategy of Conflict*. Cambridge: Harvard University Press.

[43] Schwarz, Michael and Konstantin Sonin. 2008. "A Theory of Brinkmanship, Conflicts, and Commitments." *Journal of Law, Economics, and Organization*. 24 (1): 161-183.

[44] Shapley, Lloyd. 1953. "A Value for n-person Games." In *Contributions to the Theory of Games: Volume II* eds. H.W. Kuhn and A.W. Tucker. Princeton: Princeton University Press, 307-317.

[45] Singer, J. David, Stuart Bremer, and John Stuckey. 1972. "Capability Distribution, Uncertainty, and Major Power War, 1820-1965." in Bruce Russett eds. *Peace, War, and Numbers*. Beverly Hills: Sage, 19-48. accessed through Correlates of War. National Military Capabilities (v4.0).

[46] Skaperdas, Stergios. 1998. "On the Formation of Alliances in Conflicts and Contests." *Public Choice* 96 (1): 25-42.

[47] Tooze, Adam. 2006. *The Wages of Destruction: The Making and Breaking of the Nazi Economy*. New York: Penguin Books.

[48] Wagner, R. Harrison. 1986. "The Theory of Games and the Balance of Power." *World Politics* 38 (4): 546-576.

[49] Wagner, R. Harrison. 2000. "Bargaining and War." *American Journal of Political Science* 44 (3):469-484.

[50] Wagner, R. Harrison. 2007. *War and the State: The Theory of International Politics*. Ann Arbor: The University of Michigan Press.

[51] Waltz, Kenneth. 1979. *Theory of International Politics*. Reading, MA: Addison-Wesley.

[52] Wolford, Scott. 2007. "The Turnover Trap: New Leaders, Reputation, and International Conflict." *American Journal of Political Science* 51(4): 772-788

[53] Wolford, Scott. 2012a. "Incumbents, successors, and crisis bargaining : Leadership turnover as a commitment problem." *Journal of Peace Research* 49(4): 517-530

[54] Wolford, Scott. 2012b. "Successor-Driven Wars: More Incumbents, Successors, and Crisis Bargaining." Typescript.

[55] Wolford, Scott, Dan Reiter, and Clifford Carrubba. 2011. "Information, Commitment, and War." *Journal of Conflict Resolution* 55 (4): 556-579.