

Regime Change*

Jon X. Eguia[†]
Michigan State University

September 11, 2019

Abstract

I present a theory of multi-lateral conflict. Policy-motivated countries launch military interventions in a target country, whose policies are perceived as noxious. A successful intervention leads to a change of regime and a change of policies in the target country. Comparative statics show that an intervention is more likely in a more interconnected world, if the target nation is smaller, or if the policy preferred by the target country's government is more extreme. To measure the effectiveness of alliances, I develop a measure of "relative sacrifice" in contributions to multilateral interventions. Using Afghanistan (2001-2014) as an illustration, I argue that the relative sacrifice made by the US, the UK, and Canada was high, while all other European NATO allies sacrificed very little.

*I thank Andrew Coe, Andre Little, Massimo Morelli, Debraj Ray, Arturas Rozenas, the Harvard Political Economy Reading group, and seminar audiences at NYU, U. Cyprus, U. Waterloo, U. Padova, Bocconi U., and EHU/UPV for their suggestions. Funding for this project from the European Research Council grant PCIG14-GA-2013-630097, received in 2014, is gratefully acknowledged.

[†]eguia@msu.edu.

“Regime change: the replacement of one administration or government by another, especially by means of military force.” (Google definition, May 2017).

I start with a simple motivating question: why did the United States invade Iraq in 2003? The official answer was: to remove Saddam Hussein from power. This intervention is an instance of a more general phenomenon, characterized by a military intervention by a foreign power to change the government and the policies of a weaker target nation. Other instances are the interventions in Afghanistan (2001) and Libya (2011). A common feature of these interventions is that they featured a coalition of countries supporting the intervention to change the regime in the target country. As the United States continues to mull over additional future interventions, I ask: what are the key strategic incentives that determine the decision to intervene? And what level of support -or opposition- might the US expect from other countries?¹

To help address these questions I propose a general theory of multilateral conflict applied to the phenomenon of policy-motivated regime change.

A defining feature of these US military interventions is that the goal of the attack is not to gain territory, power or wealth upon victory. Rather, the goal is to replace a ruling administration for another that would implement policies more congenial to the US. These are conflicts of regime change and policy change: a target country is invaded because its domestic policies cause a negative externality to other countries.

In a traditional divide-the-pie conflict over territory or resources, countries fight and the one that wins obtains control over the disputed resource. Third parties are affected only indirectly, insofar as the new balance of power affects the outcome of subsequent conflicts. Whereas, in a dispute over policy, all countries directly enjoy (or suffer) the policy outcome. Twenty-first century wars involving the US start because the US -and possibly its allies- want to put an end to the unwanted policy, and they do so by enacting a regime change, and installing a different administration that implements a different policy. These conflicts are better understood as a collective action problem of suppressing a collective nuisance, than

¹Potential interventions discussed in recent years include interventions in Syria (Yacoubian 2017); and in North Korea and Iran (Pompeo 2018).

as a pie-division problem with private consumption.

I model this strategic environment as a game between three or more policy-motivated countries. Countries care about their own domestic policy, and also about the domestic policy of a foreign nation that is a potential target for an intervention. Each of several “major power” countries, decides whether or not to launch an intervention against the target. All countries can invest resources to support any intervention, or to help the target country defend itself. The outcome is probabilistic, as a function of all investments. If an intervention succeeds, there is regime change in the target country, and its domestic policy is determined by the country that launched the successful intervention.

An equilibrium always exists in this game. Motivated by the interventions against the Taliban in Afghanistan in 2001, Saddam Hussein in Iraq in 2003 and Gaddafi in Libya in 2011 –and by other potential US-led interventions in the near future– I am most interested in an application with a hegemon, which is the sole major power, and in which the policy preferences of all other relevant countries are closer to those of the hegemon than to those of the target country. In this application, the target nation is a rogue country, alienated from all other countries. I identify the set of parameters for which the equilibrium outcome is peaceful, with no intervention and no regime change, and the set of parameters for which in equilibrium there is conflict. Comparative statics establish that an intervention for regime change is more likely to occur if:

- a) the world is more interconnected, so that countries are more affected from the domestic policies adopted by foreign nations;
- b) the hegemon or the major powers more intensely dislike the policies preferred by the rogue (the rogue country is more extreme);
- c) the hegemon is better at state-building, in the sense that its administration of the target country after regime change is expected to be better;
- d) the rogue country is small or the hegemon is large.

If a hegemon intervenes, smaller countries contribute in support of the intervention, but they partially free ride: smaller countries contribute less per capita than larger ones, so the

hegemon bears the lion's share of the cost of prompting regime change (and policy change) in the rogue country. For a range of parameter values, in equilibrium there is a positive probability of intervention and regime change, but an intervention is not certain to occur; rather, as the incentives to intervene increase, the rogue country invests more in its defence, keeping the hegemon indifferent about intervention.

My analysis deliberately abstains from any institutionalist considerations about the rule of law and from idealist and normative theories about a just war: I present a positive, pure rationalist theory of conflict built around the premise that amoral major powers intervene at will to suit their interests. Similarly, other countries support or oppose these interventions as it suits them according to their own interests. The basic framework is entirely non-cooperative, without commitment and without ethical constraints on behavior; I assume that countries support each other only insofar as their individual interests align. The root cause of conflict in this framework is the inability to commit to a policy compromise. My results characterize conditions under which this inability to commit to a policy leads to multi-lateral conflict aimed at regime change.²

In Section 5 I introduce alliances: how do the results change if a subset of countries are able to commit to jointly support interventions that are collectively optimal for the alliance? In theory, a committed alliances resolves the collective-action problem that leads smaller countries to free ride and to shirk in their contributions in the absence of an alliance. I test whether NATO functions as a committed alliance in practice, not just in theory. I analyze the intervention in Afghanistan from 2001 to 2014. I derive testable hypotheses on whether NATO countries contribute to the alliance efforts as if to optimize the alliance's aggregate welfare, or whether countries shirk in pursuit of their individual interests. To assess the value of the alliance, I quantify how much countries contribute to the alliance's mission, beyond their own individual interest to contribute. This measurement generates a novel "relative sacrifice" index. Using data on casualties incurred by NATO members in Iraq, I show that among the twelve largest NATO countries, the sacrifice by the US, the UK and Canada

²Lack of commitment power has long been understood to be a main cause of conflict. See for instance surveys by Fearon (1995) or Jackson and Morelli (2011).

was high, while all Continental Europe countries largely shirked, sacrificing their own self-interest very little (or not at all). The Afghanistan (2001-2014) illustration of the theory is a proof-of-concept: the theory is flexible enough to be adapted, and ultimately calibrated, to fit other applications of interest in future work.

Literature Review

Most formal research on conflict explains war as a bargaining breakdown between two parties that bargain over the division of a pie. As noted by Powell (2002), formal work on *“the origins, conduct, and termination of war [...] draws very heavily on Rubinstein’s (1982) seminal analysis of the bargaining problem.”* Rubinstein’s problem is a two-player bargaining game.³ A feature of recent military interventions is that the US builds a “coalition of the willing” or “coalition of convenience” (Kreps 2011) with other nations willing to support the intervention. To understand these conflicts, we need a multi-lateral theory of conflict. The divide-the-pie bargaining framework, so useful to study bilateral conflict, is inadequate to study multi-lateral conflict: the policy space does not resemble a pie to be consumed privately. With at least three agents and general preferences over policy, someone’s gain need not be another’s loss; rather, each policy outcome delivers a different utility profile and a different level of aggregate welfare.

In a study of military coalitions, Wolford (2015) takes a first step to relax the assumption of pure private consumption intrinsic to divide-the-pie models, by assuming that outcomes have a public-value component in addition to the private one. More generally, in the presence of ideological preferences or externalities across multiple countries, preferences over the policy space do not fit into the restrictions of either a pure public value or a pure private value. I study conflict under these more general preferences.

Traditional theories of multilateral conflict focus on power relations across countries, and in particular about “balance of power,” i.e. the idea that nations coalesce and a stable outcome is one in which two coalitions antagonistic to each other emerge, but they have

³A recent survey on foreign interventions by Aidt, Alborno and Hauk (2019) proposes a unifying framework with just two countries: a foreign power, and the target of the intervention. Chapter 7 of their survey is devoted to regime change.

similar aggregate power, so that neither coalition has an incentive to start a war (Waltz 1979; Wagner 1986; Walt 1987; Niou and Ordeshook 1990; Wagner 1994). Favorite examples refer to 1800s and pre- World War I alliances among European powers.⁴ The strategic environment leading to modern military interventions do not fit this framework: what changed from 2000 to 2001 in the relation between the US and its allies and Afghanistan and its allies (if any) was not their balance of power, which was always overwhelmingly unbalanced; rather, what changed was the US perception as to whether or not the Afghanistan represented a nuisance worth addressing through a military intervention and regime change.

My theory relates more closely to the ideas of Olson and Zeckhauser (1966) on provision to military alliances as a public good. Like them, I find that smaller countries contribute less per capita than larger ones to the military effort; unlike them, I study not just the investment decisions, but also the binary decision on whether or not to launch an intervention, and I determine the equilibrium outcome in this strategic environment.

Some more recent theories of multi-lateral conflict treat conflict as a contest for an indivisible prize (Skaperdas 1998). This framework misses out on the policy externalities that are key -I argue- to the military interventions I want to explain. Esteban and Ray (1999) enrich the contest model, allowing agents to have preferences over other groups' preferred policy outcomes. In Esteban and Ray's theory, in equilibrium all groups invest in conflict and the result of conflict is stochastic, depending on the outcome of a contest success function. This result fits their intended application to social conflict, lobbying and pressure groups well, but it's a poor prediction for an application to international relations and military interventions: not every country is always at war.

Esteban and Ray "*assume that no group expends resources on outcomes other than its preferred position.*" They acknowledge that this assumption is not entirely satisfactory. They suggest that: "*it may well be that a group decides to support the lobbying activities of some other group. A satisfactory treatment of this issue will have to depart from the present model in one of two ways. One route is to look at nonconvex lobbying technologies in which some*

⁴More recently, Bonfatti (2017) explains how geo-political struggles and trade influence foreign interventions in a world with three countries.

threshold resource expenditure is needed to influence the success probability at all.” This is the route I pursue. Substituting “intervention” for “lobbying” and “country” for “group”, I start from the premise that there is a fixed cost to merely prepare and launch a military intervention that must be incurred before the intervention has any hope of succeeding, and it then follows naturally that countries may want to support someone else’s intervention, rather than launch their own.

Bloch, Sánchez-Pagés and Soubeyran (2006), and Gallop (2017) develop other variations of Esteban and Ray’s model. Bloch et al. introduce collusion across rent-seeking groups, and their main result is that the grand-coalition is the efficient coalition structure. Gallop (2017) studies a policy dispute among three countries, and Gratton and Klose (2017) analyze conflict among three factions in a civil war. While all these theories are well suited to analyze the phenomena they study, none of them is suited to explain policy motivated, multilateral military interventions, such as those we observed in Afghanistan (2001), Iraq (2003) or Libya (2011), or potential new ones discussed since, such as Syria, Iran or North Korea.

To explain these conflicts, I propose a multilateral conflict theory of regime change.

1 The model

Overview. A set of $n+1$ countries have preferences over the policies chosen in each country. A subset of these countries are major powers that can each launch an intervention in a target country, to change its government and hence its policy. Each country can spend resources in support of any intervention by a major power, or in defending the target. A country’s welfare depends on its expenditures and on the final policy outcome. I formalize this strategic interaction as a game. I first introduce the players; then the actions and strategies they can take, with the information structure and timing of the game; and finally their preferences.

Players. Let $N \equiv \{0, 1, 2, \dots, n\}$ be a set of countries. I treat the government in each country as a unitary actor that is a strategic player in international relations, and I study the game played by these $n+1$ players. I refer to Country 0 as the “**target**” country, to the set of

countries $M \equiv \{1, \dots, m\} \subset N$ as the “**major powers**”, and if $M = \{1\}$, to Country 1 as the “**hegemon**.” For each $j \in N$, let θ_j denote the size of Country j , interpreted as a measure of power in terms of population or economic size (or a mix of both). Let $\theta \equiv (\theta_1, \theta_2, \dots, \theta_n)$ denote the row vector of sizes. Let $g \in (0, 1)$ be the (exogenously given) fraction of national resources necessary to sustain a domestic government. Then, for each $j \in N$, the government in Country j has $(1 - g)\theta_j$ resources available in the game.

Actions. Each major power simultaneously decides whether or not to launch an independent intervention into the target country. An intervention is a military attack to topple the target’s government, attempting to replace it with a new administration. For each major power $j \in M$, let $a_j \in \{0, 1\}$ denote major power j ’s decision on whether to **launch an intervention** ($a_j = 1$) or not to launch one ($a_j = 0$). Launching an intervention consumes an amount of resources that is increasing in the size of the target country. Specifically, I assume that it consumes $\theta_0 g$ resources, which represents a fraction $\frac{\theta_0 g}{\theta_j}$ of Country j ’s resources.

Simultaneously, each country decides how much to contribute to **defend the target** country. Let $d_0 \in [g, 1]$ represent the fraction of its national resources that the target country invests on setting up its government and on its defence forces. And for each $j \in N \setminus \{0\}$, let d_j denote the fraction of national resources that Country j devotes to defend the target country’s government against intervention; note that $d_j \in [0, 1 - g]$ for any $j \in N \setminus \{0\}$. Let d be the (column) vector with all the investments in the defence of the target country.

If at least one major power launches an intervention, a state of conflict arises in the target country, and all countries are called to take new actions: they can choose which fraction of their remaining resources to spend in support of each of the interventions that have been launched. Let $I \subseteq M$ be the subset of major powers that launched their own intervention. For each $j \in N$, and for each $i \in I$, let $r_{j,i} \in [0, 1 - g - d_j]$ denote the fraction of Country j ’s resources that j spends **supporting the intervention** by $i \in I$. The budget for each major power that launched an intervention, is $\sum_{h \in I} r_{i,h} \in [0, 1 - g - d_i - \frac{\theta_0}{\theta_i} g]$; the budget constraint

for any other country $j \in N \setminus I$, it is $\sum_{h \in I} r_{j,h} \in [0, 1 - g - d_j]$.⁵

For each intervening major power $i \in I$, let $r_{\cdot,i} \equiv (r_{0,i}, r_{1,i}, r_{2,i}, \dots, r_{n,i})$ denote the column vector of investment decisions in support of i 's intervention, so $\theta \cdot r_{\cdot,i} = \sum_{j \in N} \theta_{j,i} r_{j,i} \in \mathbb{R}_+$ is the total amount of resources invested in support of the intervention launched by Country i . Let r denote the matrix of all investments in support of each intervention.

Information and timing. The game has two stages: build-up stage, and conflict stage. At the beginning of the build-up stage, all parameters of the game, including the set of countries N , their vector of sizes θ , and their preferences (to be specified below) are common knowledge among all players.

1. Build-up stage - Each Country $j \in N$ chooses d_j i.e. how much to invest in defending the government of the target country. These decisions are simultaneous, and each of them is observed only by Country j and by the target country. At the same time, each major power $j \in M$ chooses $a_j \in \{0, 1\}$, i.e. whether or not to launch an intervention.

If no major power intervenes, the game ends in a peace outcome. If at least one major power launches an intervention, the set of all interventions and the total defence of the target country are publicly observed (formally, d and I are observed), and the game proceeds to the conflict stage.

2. Conflict stage - For each $j \in N$, and for each $i \in I$, Country j chooses $r_{j,i}$ i.e. how much to invest in support of intervention i . These choices are simultaneous.

At the end of the conflict stage, the conflict is resolved by the contest function (1). For each intervention $i \in I$, the probability that the intervention led by Country i succeeds is

$$\frac{\theta \cdot r_{\cdot,i}}{\sum_{h \in I} \theta \cdot r_{\cdot,h} + \theta \cdot d} \tag{1}$$

the target country successfully defends itself if all interventions fail. After conflict is resolved, the game ends in a conflict outcome and payoffs accrue, as described below.

⁵Implicit in the definition of a major power is that if $i \in M$, it is affordable for i to launch an intervention. That is, for each major power $i \in M$, I assume $\frac{\theta_0}{\theta_i} g \leq 1 - g$, or equivalently, $\theta_i \geq \theta_0 \frac{g}{(1-g)}$.

Strategies. For each major power, a strategy is a triple: a binary decision of whether to launch an intervention; a decision on the fraction of the resources spent defending the target; and, contingent on at least one intervention taking place, a decision on an allocation of resources to each intervention, as a function of the observed set of interventions and the observed set of contributions to the defence of the target.

For any other country $j \in N \setminus M$, a strategy is a pair: a decision on the fraction of resources spent defending the target; and, contingent on at least one intervention taking place, a decision on an allocation of resources to each intervention, as a function of the observed set of interventions and the observed set of contributions to the defence of the target.

Preferences. Countries have preferences over policy, and over resources.

Let P be an abstract policy space. For each $j \in N$, let $p_j \in P$ denote the policy implemented in Country j , and let $p \equiv (p_0, p_1, \dots, p_n) \in P^{n+1}$ be the profile of implemented policies. Each country is affected by the entire policy profile p . For each $j, h \in N$, let $v_{j,h} : P \rightarrow \mathbb{R}$ represent the **policy preferences** of Country j about the policy implemented in Country h . I assume that each country's policy preferences are separable across other countries' policies. In particular, assume that there exists a parameter $\phi \in (0, 1)$ and a

$(n+1) \times (n+1)$ geographic matrix $\Lambda \equiv \begin{bmatrix} 1 & \lambda_{0,1} & \cdots & \lambda_{0,n} \\ \lambda_{1,0} & 1 & & \lambda_{1,n} \\ \vdots & & \ddots & \vdots \\ \lambda_{n,0} & \lambda_{n,1} & \cdots & 1 \end{bmatrix} \in \mathbb{R}_{++}^{(n+1) \times (n+1)}$ such that

for each $j \in N$, Country j 's preferences over the policy profile $p \in P^{n+1}$ is represented by the **utility function**

$$u_j(p) \equiv v_{j,j}(p_j) + \phi \sum_{h \in N \setminus \{j\}} \lambda_{h,j} v_{j,h}(p_h).$$

Parameter ϕ represents the interconnectivity of the world, capturing technological development in transportation, travel and communication. For each pair of countries j and h , the value $\lambda_{h,j}$ captures the determinants of influence of Country h on Country j that are idiosyncratic to these pair, including their geographic distance, the terrain between them,

their cultural, ethnic and trade links and their respective sizes. Together, $\phi\lambda_{h,j}$ capture the influence or policy externality of Country h over Country j , or, conversely, how susceptible is Country j to the policy chosen in Country h .⁶

I assume that each country $j \in N$ has a unique ideal domestic policy $\hat{p}_j \in P$. I also assume that for each $j \in \{1, \dots, n\}$, the policy implemented in Country j is exogenously fixed to be equal to the country's ideal domestic policy, i.e. $p_j = \hat{p}_j$.⁷ For expositional convenience, I normalize a country's utility from their ideal domestic policy to zero, i.e. for each $j \in N$, $v_{j,j}(\hat{p}_j) = 0$.

The tension is about the policy implemented in the target country, which is determined through the strategic interaction of all countries, and resolved possibly through conflict. If no major power launches an intervention, or if all interventions fail, the target country, just like any other country, implements its ideal domestic policy, i.e. $p_0 = \hat{p}_0$.⁸ On the other hand, if an intervention by any Country $i \in I$ succeeds, then there is a change of regime, and a new administration under the tutelage of Country i takes control of the target country. I assume that the policy adopted in the target country under Country i 's administration is exogenously given, and I denote it by $p_{i,0} \in P$. I interpret $p_{i,0}$ as the best (but possibly flawed) attempt to implement Country i 's ideal policy for the target country following a change of regime led by Country i .⁹ I assume that all countries have strict preferences over $\{p_{1,0}, \dots, p_{m,0}\}$.

⁶The limit of no interconnectivity $\phi = 0$ represents the special case with no policy externalities across countries.

⁷This is consistent with citizen-candidate theories of policy-making: the policy outcome is the ideal policy of the agent who determines policy in the country, because agents cannot commit to choose a policy different from their own ideal. We could add a 3rd stage in which each country chooses policy and we would obtain the same result: without commitment, all countries would choose their ideal policy. For ease of exposition, I anticipate this outcome and I omit this third stage.

⁸The outcome in which all interventions fail represents two different real-world scenarios. The first is a literal failure of each intervention: the military attack is a flop, and the incumbent government in the target country stays in place, implementing its ideal policy \hat{p}_0 . The second scenario is one in which the initial launching of an intervention topples the existing government, and the failure of the intervention materializes at a subsequent stage of conflict for control of the country. If the government that emerges from the ashes of this conflict is aligned with the deposed incumbent and implements its ideal policy \hat{p}_0 , then the foreign intervention ultimately failed.

⁹On the principal-agent problem that an intervening major power encounters when trying to direct policy to a nominally friendly administration in a client state, see Powell (forthcoming).

With respect to the **costs**, by assumption, the fraction g of resources is already spent to set up a government before the game starts, so I consider this expenditure a sunk cost, and I normalize it to zero. Any additional investment of resources into conflict is an optional cost, as it represents a departure from the closed-economy optimum size of the government. I assume that there exists a twice continuously differentiable cost function $c : [g, 1] \rightarrow \mathbb{R}$, such that the utility cost for Country j that invests a total fraction x of its resources into its government or into conflict, either launching an intervention, supporting any intervention(s) or defending the target country, is $c(x)$. I assume that $c'(g) = 0$, $c'(x) > 0$, $c''(x) > 0$ for any $x > g$ and $\lim_{x \rightarrow 1} c'(x) \rightarrow \infty$.¹⁰

The ex-post payoff for Country j is $u_j(p) - c(x_j)$, where $u_j(p)$ is the policy utility based on the list of adopted policies p and x_j is the total expenditure of resources by Country j . The expected utility of Country j as a function of the vector of decisions to intervene $a = (a_1, \dots, a_m)$ and of the investment in resources in conflict (d, r) is

$$EU_j(a, d, r) = E_p[u_i(p|(a, d, r))] - c(x_j), \quad (2)$$

where $E_p[u_i(p|(a, d, r))]$ is the expected value of the utility over policy $u_j(p)$ given the actions (a, d, r) . The expectation is over the stochastic resolution of p_0 determined by the contest function (1).

Solution concept. Countries are strategic actors that maximize their expected utility. The solution concept is Subgame Perfect Nash Equilibrium. This completes the description of the game.

I obtain a preliminary result.

Proposition 1 *A Subgame Perfect Nash equilibrium exists.*

The interpretation of this result is that the theory is applicable, and delivers a prediction, for any set of parameter values within the scope of the model.

¹⁰These cost assumptions are microfounded on an assumption that all surplus resources not spent by the government are consumed, citizens have decreasing marginal utility of consumption, and this marginal utility is arbitrarily large as consumption is reduced to arbitrarily close to zero.

2 Application: A Rogue State

Motivated by the US-led invasion of Afghanistan in 2001, I first focus on applications in which there is a unique hegemonic major power ($M = \{1\}$) and every other country prefers that the policy implemented in the target country be the one that Country 1 (“the hegemon”) would implement there ($p_{1,0}$), to the one that the target would implement (\hat{p}_0).

Definition 1 *The target country is a rogue state if $v_{j,0}(p_{1,0}) > v_{j,0}(\hat{p}_0)$ for any $j \in N \setminus \{0\}$.*

For any $n \in \mathbb{N}$ such that $n \geq 2$, let \mathcal{G}^n denote the class of games with a rogue state, a hegemon, and $n - 1$ other countries. Let $\mathcal{G} \equiv \bigcup_{n=2}^{\infty} \mathcal{G}^n$, and let $\Gamma \in \mathcal{G}$ denote an arbitrary game with a rogue state, a hegemon, and $n - 1$ other countries. In the running example in mind, the rogue state is the Taliban regime in Afghanistan in 2001, the hegemon is the United States, and the other $n - 1$ countries are the US western allies.

Definition 2 *A peace equilibrium is a pure strategy equilibrium with no intervention. An intervention equilibrium is a pure strategy equilibrium with an intervention.*

I show in the Appendix that for any $z \in [g, 1]$, the conflict subgame that follows after the rogue chooses $d_0 = z$, for each $j \in N \setminus \{0\}$ Country j chooses $d_j = 0$, and the hegemon chooses to intervene, has a unique pure strategy equilibrium (Lemma 15).¹¹ For each $j \in N \setminus \{0\}$, and for any $d_0 \in [g, 1]$, let $r_j(d_0)$ be the equilibrium investment by Country j in support of the intervention in this subgame, as a function of d_0 . Let $r(d_0)$ denote the column vector of such investments. Note the first component of $r(d_0)$ is $r_0(d_0) = 0$.

Lemma 2 *For any $\Gamma \in \mathcal{G}$, a peace equilibrium exists if and only if*

$$\underbrace{\phi \lambda_{0,1} (v_{1,0}(p_{1,0}) - v_{1,0}(\hat{p}_0))}_{\text{Policy gain if success}} \underbrace{\frac{\theta \cdot r(g)}{\theta \cdot r(g) + \theta_0 g}}_{\text{Prob. of success}} \leq c \underbrace{\left(g + \frac{\theta_0}{\theta_1} g + r_1(g) \right)}_{\text{Hegemon's cost}}. \quad (3)$$

¹¹Notice that the probability that an intervention succeeds increases in investments in support of an intervention, but the derivative of this probability decreases in resources spent: the marginal return from investment is positive but decreasing. The derivative of the probability that the intervention succeeds with respect to investments in support of the intervention is first increasing and then decreasing in investments in defence of the rogue (the cross-partial derivative is single-peaked, with a peak at the point that investments in support equal investments in defence).

Furthermore, if this condition is satisfied, this equilibrium is unique.

Notice that the peace equilibrium condition (3) holds if:

- a) the interconnectivity parameter ϕ is sufficiently small (countries are not much affected by the domestic policies of other nations);
- b) the proximity parameter $\lambda_{0,1}$ is sufficiently small (the rogue is far from the hegemon);
- c) $v_{1,0}(\hat{p}_0)$ is sufficiently large (the rogue nation is not too rogue);
- d) $v_{1,0}(p_{1,0})$ is sufficiently small, i.e. negative and of high magnitude (the hegemon is bad at foreign state-building);
- e) θ_0 is sufficiently large (the rogue nation is too big to be bullied);
- f) θ_1 is sufficiently small (the hegemon is weak); or
- g) g is sufficiently large (launching an intervention is very costly).

Given an arbitrary game $\Gamma \in \mathcal{G}^n$ with a hegemon, a rogue nation, and $n - 2$ other nations, consider a simplified game $\tilde{\Gamma}$ of investments in conflict played by the same set of nations, in which an intervention is fixed exogenously to occur ($a_1 = 1$ as a non-strategic variable), the rogue nation decides how much to invest to defend its own domestic government, no other country supports the rogue, and observing the rogue's investment, all other countries decide how much to invest to support the intervention.

I show in the Appendix that game $\tilde{\Gamma}$ has a pure strategy equilibrium (Lemma 16). Let \tilde{d}_0 be the rogue country's equilibrium investment on its own government in game $\tilde{\Gamma}$. If the equilibrium is non-unique (a knife-edge case),¹² let this notation refer to the equilibrium with the lowest value of \tilde{d}_0 .

With this notation at hand, return to the full game $\Gamma \in \mathcal{G}^n$ in which the hegemon decides strategically whether or not to intervene in the rogue country, and in which all other countries strategically choose their investments in support of the intervention (or in support of the rogue).

¹²More formally, consider a set of cost functions characterized by a finite set of parameters; then uniqueness is generic in parameter values.

Lemma 3 For any $\Gamma \in \mathcal{G}$, an intervention equilibrium exists if and only if

$$\phi \lambda_{0,1} (v_{1,0}(p_{1,0}) - v_{1,0}(\hat{p}_0)) \frac{\theta \cdot r(\tilde{d}_0)}{\theta \cdot r(\tilde{d}_0) + \theta_0 \tilde{d}_0} \geq c \left(g + \frac{\theta_0}{\theta_1} g + r_1(\tilde{d}_0) \right). \quad (4)$$

If this inequality does not hold, the hegemon prefers to not intervene, if its intervention is anticipated. If the hegemon prefers to intervene if its intervention is not anticipated and it prefers to not intervene if its intervention is anticipated, then there is no equilibrium in pure strategies, and the equilibrium must be in mixed strategies.

Lemma 4 For any $\Gamma \in \mathcal{G}$, if neither (3) nor (4) holds, there exists a mixed strategy equilibrium in which the hegemon intervenes with probability $\sigma \in (0, 1)$ and $d_0 \in (g, \tilde{d}_0)$.

In this mixed strategy equilibrium, the rogue hedges: expecting an intervention to take place only with some probability, it invests in defence more than if it were certain to face no intervention, but less than if an intervention were certain to take place. In fact, it invests in defence just enough to leave the hegemon indifferent about launching an intervention.

Putting together lemmas 2, 3 and 4, we infer that if the connectedness parameter ϕ , the proximity parameter $\lambda_{0,1}$ or the policy improvement from intervening $v_{1,0}(p_{1,0}) - v_{1,0}(\hat{p}_0)$ are low (i.e. if the hegemon has little to gain from an intervention), or if $\frac{\theta_0}{\theta_1}$ or g are high (an intervention is costly and less likely to succeed), the unique equilibrium is such that there is no conflict. As incentives for conflict increase or the difficulty of succeeding in conflict decreases, this peaceful equilibrium breaks down, and we face uncertainty with a positive probability of intervention in a mixed strategy equilibrium; the rogue nation adjusts accordingly, increasing its allocation of resources to defend its government. As the incentives for conflict or the ease of success in intervention increase, this mixed equilibrium involves greater resources devoted to conflict and greater probability of intervention: it thus moves continuously from the peace equilibrium to the intervention equilibrium. If the incentives for conflict are sufficiently large, or the resources that an intervention needs to be likely to succeed are sufficiently small, then the equilibrium is sure to involve an intervention and positive investments in conflict by all countries.

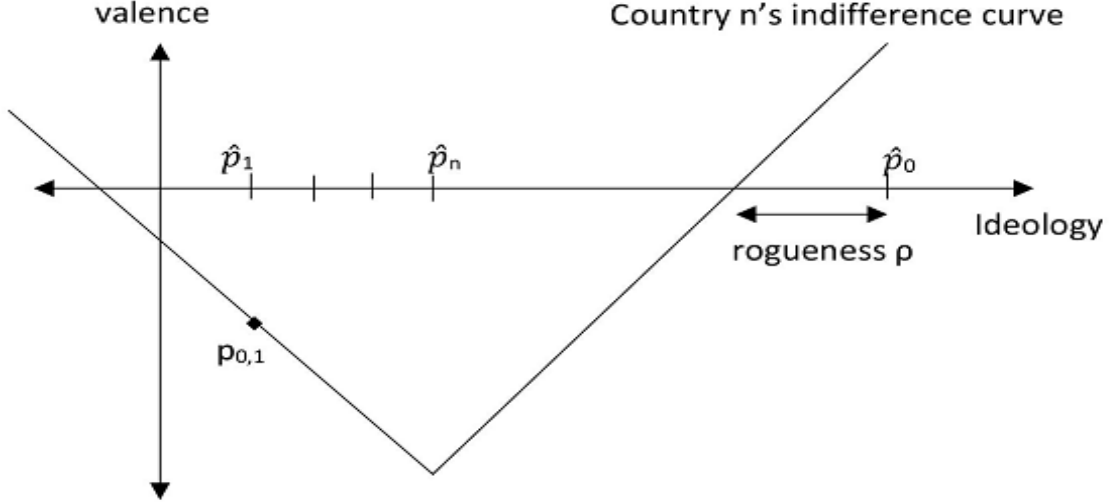


Figure 1: A rogue in a space with preferences over ideology and valence

I formalize the first of these comparative statics, with respect to the connectedness parameter ϕ . Informally, if ϕ is low, other countries are not affected much by the policy of the rogue nation, so there is no motivation to intervene; whereas, if ϕ is large, the domestic policy of the rogue country is very relevant and the incentive to intervene abroad is greater.

Proposition 5 *For any $\Gamma \in \mathcal{G}$, there exist $\phi_{Peace} \in (0, 1]$ such that*

- (i) *if $\phi \in [0, \phi_{Peace}]$, the unique equilibrium is a peace equilibrium, and*
- (ii) *if $\phi > \phi_{Peace}$, an intervention occurs with strictly positive probability.*

Under appropriate restrictions on admissible preferences, we can draw a comparative static relating the extremism or “rogueeness”, of the rogue country, to the chances of an intervention. To quantify “extremism”, assume –only for this result– that the policy space is two dimensional, with an ideological first dimension over which each country has Euclidean preferences around its ideal, and an additively separable common-value second dimension that captures valence or quality.

Assume that on the ideological first dimension, for each $j \in N$, Country j 's ideal domestic policy \hat{p}_j is also Country j 's ideal for the policy to be implemented in the rogue country, and assume that the hegemon and the rogue country have ideal policies at opposite extremes of the ideological spectrum. Without (further) loss of generality, relabel countries so that $\hat{p}_1 \leq_I \hat{p}_2 \leq_I \dots \leq_I \hat{p}_{n-1} \leq_I \hat{p}_n <_I \hat{p}_0$, where \leq_I is the left-to-right order in the ideological dimension, and assume $p_{1,0} = \hat{p}_1$, that is, the hegemon would impose its own ideal ideological policy on the target following regime change. Further, assume that on the valence dimension, the valence of the target's ideal domestic policy is zero, whereas the valence of policy $p_{1,0}$ is $\mu < 0$, as in Figure 1. Then, for any $j \in N$, $v_{j,0}(p_{1,0}) - v_{j,0}(\hat{p}_0) = -(\hat{p}_j - p_{1,0}) + \mu + (\hat{p}_0 - \hat{p}_j)$ and the target country is rogue if and only if $-(\hat{p}_n - p_{1,0}) + \mu + (\hat{p}_0 - \hat{p}_n) \geq 0$, that is, $\hat{p}_0 \geq 2\hat{p}_n - p_{1,0} - \mu$. Define $\rho \equiv \hat{p}_0 - (2\hat{p}_n - p_{1,0} - \mu)$, so that ρ measures the rogueness: $\rho = 0$ means that the country most friendly to the target is indifferent between the target's preferred policies and the ones the hegemon would implement in the target after regime change. Let $\mathcal{G}^\rho \subset \mathcal{G}$ denote the collection of games with these two-dimensional (ideological and valence) preferences, and consider a comparative static over ρ .

Proposition 6 *For any game $\Gamma \in \mathcal{G}^\rho$, there exist $\rho_{Peace} \in \mathbb{R}$ and $\rho_{Conflict} \in \mathbb{R}_+$ such that $\rho_{Peace} < \rho_{Conflict}$ and such that*

- (i) *if $\rho \in [0, \rho_{Peace}]$, the unique equilibrium is a peace equilibrium, and*
- (ii) *if $\rho \in (\rho_{Peace}, \rho_{Conflict})$, an intervention occurs with strictly positive probability, and*
- (iii) *if $\rho > \rho_{Conflict}$, the unique equilibrium is an intervention equilibrium.*

The intuition and proof of this result are similar to those of Proposition 5: the rogue nation can become a more pressing problem either because the world becomes more interconnected and thus other countries become more susceptible to the rogue's policies (ϕ higher), or because the rogue nation becomes intrinsically more problematic by adopting more noxious policies (ρ high). Either way, an intervention becomes more likely.

Alternatively, consider the effect of country size or power: interventions become less likely, and eventually stop altogether, as the rogue country becomes bigger so that changing

its regime becomes more difficult. Consider any arbitrary vector of country sizes such that the hegemon is the largest country. For each $j \in N \setminus \{0\}$, define $\tilde{\theta}_j \equiv \frac{\theta_j}{\theta_1}$. Fix a vector $\tilde{\theta} = (1, \tilde{\theta}_2, \tilde{\theta}_3, \dots, \tilde{\theta}_n) \in [0, 1]^n$ of relative sizes for all countries but the rogue, such that $\sum_{j \in N \setminus \{0\}} \tilde{\theta}_j = 1$ and consider the class of games $\mathcal{G}^{\tilde{\theta}}$ such that $\theta_j = \tilde{\theta}_j(1 - \theta_0)$ for each $j \in N \setminus \{0\}$, where the degree of freedom is the size of the rogue nation, θ_0 . We can now study the comparative statics with respect to the rogue's size.

Proposition 7 *Consider all games in $\mathcal{G}^{\tilde{\theta}}$. There exist $\theta_0^{Peace} \in (0, 1)$ and $\theta_0^{Conflict} \in (0, 1)$ such that $\theta_0^{Conflict} < \theta_0^{Peace}$ and such that*

- (i) *if $\theta_0 \in (0, \theta_0^{Conflict}]$, an intervention occurs with certainty, and as θ_0 vanishes, the probability of regime change converges to one,*
- (ii) *if $\theta_0 \in (\theta_0^{Conflict}, \theta_0^{Peace})$, an intervention occurs with strictly positive probability, and*
- (iii) *if $\theta_0 > \theta_0^{Peace}$, the unique equilibrium is a peace equilibrium.*

The proof is again similar, but the intuition is a bit different: as the size of the rogue changes, it's not so much that the benefits of an intervention change, as in the two previous results, but rather, the benefit stays constant, while the cost of intervening increases with the size of the rogue, so if the rogue is too large, interventions are not worthwhile.

3 Multiple Interventions

I now return to the general case, in which two or more major powers have the ability to launch independent interventions and the target country may not be a rogue. This general case encompasses applications to a bipolar world with two antagonistic major powers, such as the USA and the USSR during the Cold War (1945-1989).

The qualitative results for a world with a hegemon generalize, with some qualifications.

Lemma 2 applies, but the peace condition must hold for each of the major powers independently. For each $i \in M$, and for any $d_0 \in [g, 1]$, let $r_{.,i}(d_0, \{i\})$ denote the vector of investments in support of the intervention by major power i , in the conflict subgame that follows after the target chooses d_0 , all other countries choose $d_i = 0$, and only i intervenes.

Lemma 8 *A peace equilibrium exists if and only if*

$$\phi \lambda_{0,i} |v_{i,0}(p_{0,i}) - v_{i,0}(\hat{p}_0)| \frac{\theta \cdot r_{\cdot,i}(d_0, \{i\})}{\theta \cdot r_{\cdot,i}(d_0, \{i\}) + \theta_0 g} \leq c \left(g + \frac{\theta_0}{\theta_j} g + r_{i,i}(g) \right) \text{ for each } i \in M. \quad (5)$$

Furthermore, if this condition is satisfied and the target country is a rogue, the peace equilibrium is unique.

If the target country is not a rogue, a second equilibrium can emerge. Suppose there are two major powers 1 and 2, opposed to each other. Suppose $v_{i,0}(\hat{p}_i) - v_{i,0}(p_{0,i})$ is close to zero but $v_{i,0}(\hat{p}_i) - v_{i,0}(p_{0,j})$ is very large for $\{i, j\} = \{1, 2\}$. Then aside from the (Pareto superior) peace equilibrium, we also have a conflict equilibrium in which the two major powers intervene to thwart each other's intervention.¹³

The comparative static of Proposition 5 generalizes. Let $\Gamma^{M,\phi}$ denote the game with a set M of major powers, and with a connectivity parameter ϕ .

Proposition 9 *There exist $\phi_{Peace}^M \in (0, 1]$ and ϕ_{Unique}^M such that*

- (i) if $\phi \in [0, \phi_{Peace}^M]$, a peace equilibrium of game $\Gamma^{M,\phi}$ exists and if $\phi \in [0, \phi_{Unique}^M]$ this equilibrium is unique, and*
- (ii) if $\phi \geq \phi_{Peace}^M$, at least one intervention occurs with strictly positive probability.*

In a world with a hegemon, the ideal domestic policy of the target country determines whether it is intervened or not (Proposition 6). With multiple major powers, the incentives to intervene depend not only on the ideal domestic policy of the target, but also on the policies that other major powers would implement in the target if they intervene and on their decisions to intervene, and Proposition 6 no longer applies. The comparative static with respect to the size of the target is now subject to the same complexity as the comparative static with respect to the connectivity parameter ϕ . If the target is too large, no major power

¹³The invasions of Norway by both the UK and Nazi Germany in the same week of April 1940 are illustrative of the incentives for dual interventions. The Norwegian status as a non-belligerent was satisfactory to both the UK and Nazi Germany, but both feared that the other power would intervene to alter the terms of Norwegian neutrality, so they both intervened to attempt to forestall each other's intervention. The invasion of Iceland by the UK a month later had a similar motivation.

intervenes; however, even if the target is very small, intervening is not always a dominant strategy (as it was for a hegemon): if other major powers launch their own interventions, an additional one by major power j can make the expected outcome worse for j , dissuading j from intervening. No matter how small the target, it is then possible that multiple major powers intervene only probabilistically, so that it can be that ex-post there is no intervention.

4 Alliances

Let a commitment alliance be an institutional agreement by a subset of countries $A \subset N$ (the “Alliance”) to collude to choose their actions with the goal of maximizing the aggregate payoff to all countries in the alliance, weighed by the size of the alliance members.¹⁴ Motivating examples include the old Commonwealth with its Dominions; or the defensive aspect of NATO. Assume that policy preferences are such that the target country is a rogue country. Assume there is a hegemon that belongs to the Alliance.

Let Γ^A be a game with a commitment alliance among A such that the hegemon belongs to A , and the rogue country does not ($1 \in A$, $0 \notin A$). For any $d_0 \in [g, 1]$, and for each $i \in N \setminus \{0\}$, let $r_i^A(d_0)$ be the equilibrium investments by Country i in support of the intervention in the subgame after the rogue invests d_0 in its own defence, no country helps to defend the rogue nation, and the alliance intervenes. Let $r^A(d_0)$ denote the vector $(0, r_1^A(d_0), r_2^A(d_0), \dots, r_n^A(d_0))$. Slightly adapting Lemma 2, we obtain that a peace equilibrium exists if and only if

$$\begin{aligned} & \phi \left(\sum_{i \in A} \frac{\theta_i \lambda_{0,i} (v_{i,0}(p_{1,0}) - v_{i,0}(\hat{p}_0))}{\sum_{k \in A} \theta_k} \right) \frac{\theta \cdot r^A(g)}{\theta \cdot r^A(g) + \theta_0 g} \\ & \leq \frac{\theta_1}{\sum_{k \in A} \theta_k} c \left(g + \frac{\theta_0}{\theta_1} g + r_1^A(g) \right) + \sum_{i \in A \setminus \{1\}} \frac{\theta_i}{\sum_{k \in A} \theta_k} c \left(g + r_i^A(g) \right). \end{aligned} \quad (6)$$

Furthermore, if this condition is satisfied, this equilibrium is unique. The proof replicates

¹⁴This is a stylized version of an alliance with full commitment. For a more nuanced study of the different classes of military coalitions, see Benson (2012), Wolford (2015), and the taxonomy of alliances by Benson and Clinton (2016).

the proof of Lemma 2, treating the alliance A as a single player.

We can also slightly adapt Lemma 3 to obtain a necessary and sufficient condition for an intervention equilibrium to exist. The condition is as in Lemma 3, with the exception that benefits and costs are calculated for the average over all alliance members, instead of just the hegemon.

With a commitment alliance, we obtain a qualitatively similar result on whether interventions occur or not. If the hegemon counts on the support of its committed allies, then it is more likely to launch an intervention against a rogue state if:

- a) countries are more interconnected (ϕ is higher),
- b) the rogue country is more rogue ($v_{i,0}(\hat{p}_0)$ decreases for each $i \in A$),
- c) the hegemon is better at state building ($v_{i,0}(p_{1,0})$ increases for each $i \in A$), or
- d) the rogue country is smaller (θ_0 small).

We can formalize these intuitions. With respect to the connectedness parameter, the result, as a corollary of Proposition 5, is as follows.

Corollary 10 *In game Γ^A , there exist $\phi_{Peace}^A \in \mathbb{R}_{++}$ such that*

- (i) *if $\phi \leq \phi_{Peace}^A$, the unique equilibrium of game Γ^A is a peace equilibrium, and*
- (ii) *if $\phi > \phi_{Peace}^A$, an intervention occurs with strictly positive probability.*

The exact cutoffs for intervention change as a result of the alliance. With an alliance, if there is an intervention, members of the alliance increase their investment of resources into the conflict, while non-members reduce them. The rogue country becomes worse off: it either invests more on defence, or interventions are more likely to succeed (or both). If the alliance is sufficiently large, encompassing all other countries, then an intervention becomes more likely (the peace cutoff unambiguously becomes smaller).

Proposition 11 *Assume a set of countries A with $\{1, i\} \subseteq A \subseteq N \setminus \{0\}$ for some $i \in N \setminus \{1, 0\}$ forms a commitment alliance. Then, compared to the benchmark with no alliance,*

- i) if an intervention occurs, alliance members increase their contributions to support it, and non-members decrease them,*

- ii) for parameters such that an intervention occurs with positive probability, either the rogue country invests more on its own defence, or interventions are more likely to succeed (or both), and*
- iii) if $A = N \setminus \{0\}$, the set of parameters for which an intervention occurs with positive probability expands ($\phi_{Peace}^A < \phi_{Peace}$).*

Alliance members increase their contributions because they now internalize the positive externality to other alliance members. Non-members contribute less because of the substitution effect on contributions (contributions by other countries reduce the incentive to contribute). In the aggregate, the positive effect of increased contributions dominates, so either the rogue country defends itself more, or an intervention is more likely to lead to regime change. If all countries ally against the rogue, interventions become more likely because an alliance guarantees that interventions are efficiently supported, while without an alliance they are not. This result does not extend to cases in which some countries are not members of the alliance, because (by Proposition 11 part (i)), forming an alliance shifts the burden of supporting the intervention from non-members to members. For some parameters, it is not worthwhile for a small alliance to launch an intervention and to bear its burden mostly on its own, even though it would be worthwhile if the costs were to be more widely distributed among all countries.¹⁵

Corollary 10 and Proposition 11 show that the formation of a committed alliance affects the equilibrium outcomes. A question arises: are commitments to an alliance upheld in practice? Do countries in an alliance behave as if committed to maximize the alliance's joint welfare? Or do they pay lip service to the alliance, and continue to optimize according to their own independent interests? In the next subsection, I test whether NATO functioned like a committed alliance, or like a loose umbrella of somewhat like-minded countries pursuing their own independent interests, during the 2001-14 intervention in Afghanistan.

¹⁵This result leads to the following intriguing implication for the future of NATO: if the relative power (in share of world population and/or GDP) of the countries currently in NATO continues to decrease, the current NATO alliance may no longer incentivize collective action. Will a NATO alliance of such ineffectively small size choose to either dissolve or expand to better share with other countries the costs of interventions?

4.1 Case Study: Afghanistan (2001)

I revisit the first of the US-led invasions that motivated this theory: the intervention launched in October 2001 to topple the Taliban regime in Afghanistan. After the Taliban government fell in November 2001 and Hamid Karzai was selected as interim President at a conference in Bonn in December 2001, the occupying forces were reorganized into the NATO-led International Security Assistance Force (ISAF), and the resolution 1386 of the UN Security Council tasked ISAF with maintaining security around the seat of government (Kabul).¹⁶ All NATO nations contributed troops to ISAF, as did 22 non-NATO nations. At peak deployment in 2012, ISAF numbered 132,000 troops. ISAF ceased operations and was disbanded in 2014. At the time of disbanding ISAF, the intervening nations had suffered over 3,400 deadly casualties in Afghanistan.¹⁷

I analyze this multilateral intervention with the aid of the theory. I formalize the intuitions of Olson and Zeckhauser’s (1966) collective action model, using the theoretical framework of the previous sections. A key premise is that countries choose an optimal investment in support of an intervention by equating expected marginal policy gains with marginal costs. If NATO countries were committed to maximize their joint welfare, they would minimize their aggregate cost for any investment, which –because each country’s cost function is strictly convex– implies that all members of the alliance would incur the same marginal cost.

Whereas, if NATO countries acted as independent decision-makers, maximizing their own individual interest, then each country would equate its marginal benefit to its marginal cost. Specifically, if countries optimize independently, once an intervention is launched, at the conflict stage, each country $i \in N \setminus \{0, 1\}$ chooses a contribution of resources r_i that solves

$$\phi \lambda_{0,i}(v_{i,0}(p_{1,0}) - v_{i,0}(\hat{p}_0)) \frac{\theta_i \theta_0 d_0}{\sum_{k=1}^n \theta_k r_k - \theta_0 d_0} = c'(g + r_i).$$

Comparing any two countries $i, j \in N \setminus \{0, 1\}$, in equilibrium the ratio of their marginal

¹⁶Subsequent resolutions expanded this role to maintaining security throughout the country.

¹⁷Source: icasualties.org. The ISAF mission did not report casualty totals, referring instead to its members to each report its own country-specific casualty tallies.

costs is equal to

$$\frac{c'(g + r_i)}{c'(g + r_j)} = \frac{\phi\lambda_{0,i}(v_{i,0}(p_{1,0}) - v_{i,0}(\hat{p}_0))\theta_i}{\phi\lambda_{0,j}(v_{j,0}(p_{1,0}) - v_{j,0}(\hat{p}_0))\theta_j} \quad (7)$$

For any two NATO members i and j that share the same preference for the policies that a pro-Western government would implement in Afghanistan, over the policies that a Taliban government would implement, the ratio of marginal costs incurred by members i and j under individual optimization must be equal to

$$\frac{\lambda_{0,i}\theta_i}{\lambda_{0,j}\theta_j} \quad (8)$$

If we are willing to make any assumption on $\frac{\lambda_{0,i}}{\lambda_{0,j}}$, the hypothesis of individual optimization becomes testable, subject to this assumption. Assume that $\lambda_{0,i} = \lambda_{0,j}$ for any $i, j \in N \setminus \{0\}$, that is, countries i and j are similarly affected by Afghanistan's policies.

I use the population of Country i as its size.¹⁸ I use the number of fatal casualties as the resources that each nation incurred in support of the intervention. Data on monetary expenditures for participation in ISAF for each country is difficult to obtain, and resources in the form of human lives spent in support of the intervention are not only important in itself, but also serve as proxy for the monetary cost. So I use the casualty data of Country i , and I divide it by Country i 's population to obtain the per capita casualty data that I use as a proxy for the fraction of national resources r_i that each NATO member i invested in support of the ISAF mission.

The cost to nation i of investing r_i is $c(r_i)$. Because the contributions to ISAF represented only a very small fraction of national resources, we do not need to know the exact shape of the cost function c .¹⁹ The cost in lives and money, while huge in absolute terms, it is sufficiently small as a fraction of total resources, that we are only interested in a very small

¹⁸Results using GDP as the measure of size are available from the author. Since GDP per capita is similar among the US and its Western NATO allies, results look similar for these countries; whereas, poorer Eastern European NATO allies contributed much more as a ratio of GDP, than as a ratio of population.

¹⁹For instance, the US lost on average from 2001 to 2015 less than one millionth of its population each year in operations in Afghanistan, and the Congressional Research Service estimates that the total cost of these operations was \$686 billion (Belasco 2014), which is about 0.3% of GDP in this period.

subset of the domain of c , and hence, we can approximate c by its quadratic approximation \tilde{c} . The derivative \tilde{c}' of this approximation is an affine function with $\tilde{c}'(g + r_i) = \kappa r_i$ for some $\kappa \in \mathbb{R}$, so for any two countries i and j , the ratio of their marginal approximated costs is

$$\frac{r_i}{r_j} \tag{9}$$

Therefore, since marginal benefits (Expression 8) must be equal to marginal costs (Expression 9), the theoretical prediction for any two countries $\{i, j\}$ with similar preferences is that under individual optimization,

$$\frac{r_i}{r_j} = \frac{\theta_i}{\theta_j}.$$

Hypothesis 1: *If NATO countries optimize their individual interests without collusion, then for any two members with similar preferences, the ratio of their per capita casualties is similar to the ratio of their respective populations.*

Note that this means that small countries contribute less per capita, not just in absolute numbers.

If, on the other hand, NATO countries collude to maximize the aggregate welfare of the alliance, then their per capita casualties would be similar: they would all contribute in proportion to the population: $\frac{r_i}{r_j} = 1$.

Hypothesis 2: *If NATO countries optimize their joint interest to maximize the aggregate welfare of the alliance, then for any two countries in the alliance, the ratio of their per capita casualties would be close to one.*

We can verify which of these hypothesis better reflects the actual contributions of NATO members to ISAF. The following figure summarizes the findings. I consider only the twelve largest NATO countries, to avoid the noise that comes with the small numbers of casualties suffered by small countries, and I drop Turkey, which being a predominantly Muslim country, did not have policy preferences entirely aligned with those of the rest of NATO countries (and it did not allow its troops in ISAF to engage in combat).

The first four columns of Table 1 lists the largest NATO countries,²⁰ their 2010 population in millions, their total casualties from the launch of the intervention in October 2001 to the end of the ISAF mission in December 2014, and their casualties per million people.

	(2) Pop.	(3) Casualties	(4)=(3)/(2)	(5) Ind. Optimum	(6) Sacrifice
US	318	2271	7.14	2.90	100 %
Germany	82	57	0.70	0.75	-0.1%
France	66	88	1.33	0.60	11.0%
UK	65	453	6.97	0.59	97.4%
Italy	60	53	0.88	0.55	5.0%
Spain	46	35	0.76	0.42	5.0%
Poland	39	44	1.13	0.36	11.4%
Canada	35	158	4.51	0.32	61.5%
Benelux	28	26	0.93	0.26	9.8%
Romania	19	23	1.21	0.17	21.5%

Table 1: Contributions of NATO members to the war in Afghanistan (2001-14).

We see that countries suffered very different casualty rates. How should we interpret this data? According to our theory, we should compare the ratios of per capita casualties (column (4)) to the ratios of population.

We cannot say whether NATO as a whole invested too much or too little in Afghanistan, because we do not have a reliable estimate of the magnitude of benefits and costs. What we can do is to compare countries' contributions to each other. Security in Afghanistan was a public good for NATO. A general result on public good provision is that individual contributions are inefficiently low. Hence, it is safe to assume that no country contributed more than the amount that maximized aggregate welfare. Therefore, the highest (per capita) contribution observed in the data is a lower bound on the collectively optimal contribution.

We can then compute a lower bound on the individually optimal contribution for each country, by multiplying this highest individual country contribution (7.14 casualties per million inhabitants) by the country's fraction of the total population. That's Column 5. Hypothesis 1 is that countries contributed in proportion to this amount, i.e. that Column 4

²⁰I treat the BeNeLux as a single unit of observation, because of their intense cooperation and the partial integration of their armed forces (see, for instance, their 2012 Benelux Declaration on Defence Cooperation). Independently, each of the three would be too small, best dropped.

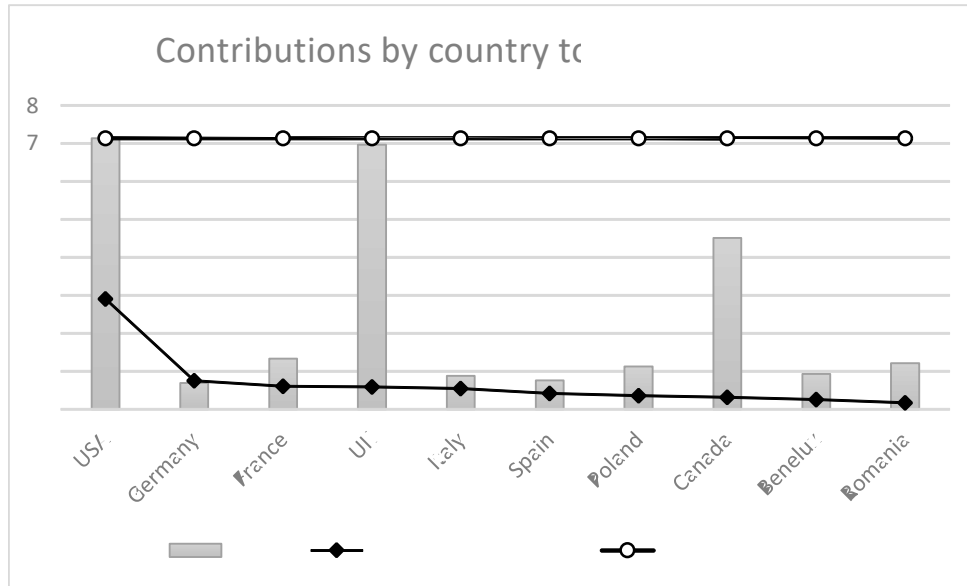


Figure 2: Casualties per capita in Afghanistan 2001-14, by intervening country.

is a factor weakly greater than one of Column 5. This exercise is inexact, in that it ignores all the other small NATO members, but it serves as a first approximation. Hypothesis 2 is that they all contributed the same.

Figure 2 illustrates our findings. The top line is the lower bound on the NATO-optimal contribution that each country should have made to maximize their joint welfare. This top line is the prediction by Hypothesis 2: equal contributions by all countries.

The bottom line is the lower bound on the optimal contributions assuming that countries maximize their individual welfare. This is the prediction by Hypothesis 1. If the collective and individual optima are in fact strictly greater than their respective lower bounds, the predictions by both hypotheses must be scaled up proportionally, but the relative magnitudes across countries or hypotheses is unchanged.

The bars indicate the actual results. They show that the US, the UK and to some extent Canada contributed as if to maximize aggregate welfare for NATO (or, at least, contributed close to the lower bound on the optimal contribution), whereas all continental EU nations (Germany, France, Italy, Spain, Poland, the Benelux and Romania) contributed little, in

line with their own individual incentives.²¹ The Pearson correlation coefficient r between country size and per capita casualties is 0.39, lending support to Hypothesis 1.

For each country, I compute a measure of **relative sacrifice**: the fraction it contributed of the difference between the (lower bound on) the collectively optimal contribution, and the country's individually optimal contribution. This statistic is a measure of the country's relative sacrifice for the alliance: it does not credit countries for what they would do anyway out of their own individual interest, and it only takes into account the additional contributions they make for the sake of the alliance's aggregate welfare. This is Column (6) in Table 1.

According to this measure, the US and the UK fulfilled its duty to the alliance fully. Canada, mostly. Most others sacrificed very little, or close to not at all.

Two observations:

One – Germany, according to this measure, contributed exactly its own selfish interest. I conjecture that this is a coincidence: Germany's military and foreign policy remain idiosyncratic, as a legacy from World War II.²²

Two – Because small countries are expected to contribute less (per capita) than bigger ones, for a given level of casualties per capita, the measure of relative sacrifice is higher for the smaller country. For instance, France suffered more casualties per capita than Romania, but given that France is three times as populous, it was expected to suffer three times as many per capita casualties if they both followed individual incentives. Hence, despite suffering lower per capita casualties, Romania incurred twice as great a relative sacrifice than France.

5 Discussion

I have presented a theory of multilateral conflict in which countries launch military interventions in a target nation. A successful intervention replaces the government of the

²¹If the actual optimum was strictly higher than the lower bound, all countries contributed below the collective optimum. However, if the actual optimum was over 60% higher than the bound, then countries such as Italy contributed below their individual self-interest optimum, which appears unlikely.

²²We could model this German idiosyncrasy with a German-specific cost function that was steeper than the general one.

target country with a new government that implements policies preferred by the nation that launched the intervention.

Motivated by the experiences of US-led interventions in Afghanistan to topple the Taliban (2001), in Iraq to topple Saddam Hussein (2003), or in Libya to topple Gaddafi (2011), I focus on the case in which a hegemon (such as the United States) decides whether or not to intervene in a rogue country, while a number of smaller countries choose how much to support the intervention, conforming a “coalition of the willing.”²³

The incentives to intervene, and to support an intervention, are driven by the externalities that the domestic policy of the rogue country exerts on other countries. The aim of the intervention is to change the government of the rogue country. Because a successful intervention benefits not just the country that launches it, but also all other countries with similar policy preferences, countries face a collective action problem: the hegemon bears a disproportionate amount of the cost, and the smaller nations partially free-ride on its efforts.

An intervention occurs if the disutility of the domestic policy position of the rogue nation is sufficiently large (if the rogue nation is too rogue); or, for a fixed policy position, if the susceptibility of countries to the domestic policy of foreign nations is sufficiently large, or if size of hegemon relative to the rogue nation is sufficiently large. These results can help us reinterpret “Democratic Peace Theory” (democracies do not fight each other)²⁴: democracies do not fight each other because, more generally, countries with similar preferences do not fight each other, and democracies have similar preferences.

Results partially generalize to an environment in which several countries can launch interventions.

If countries form an alliance, then an intervention is more likely to occur and members of the alliance invest more in support of the intervention.

As a case study, I analyze the intervention in Afghanistan in 2001-14 through the lense of

²³President Clinton used this term to refer to a possible coalition to attack North Korea in 1994, and President Bush’s administration popularized it in reference to the countries supporting the 2003 intervention in Iraq. See the White House release from June 5, 1994, “Interview with the President” by Sam Donaldson ABC; and Steve Schifferes, “US says ‘coalition of willing’ grows,” BBC News US Edition, 21 March 2003.

²⁴See for instance Bueno de Mesquita, Morrow, Siverson and Smith (1999).

the theory. Under the assumption that all Western (non-Muslim) NATO countries benefited equally from replacing the Taliban in Afghanistan, I calculate the individual incentives for each country to contribute to the intervention, and I develop and compute a measure of “sacrifice” as the ratio of the additional contributions that each country made above this individual optimum. According to this measure, the US, the UK and Canada sacrificed more, while Continental European countries sacrificed -relatively- very little.

The framework is flexible and can be adapted to fit many other applications. In any application of the theory to a specific prospective intervention, we can slightly tweak the model by relaxing some simplifying assumptions, and we can adopt specific parameter values to better fit the application under consideration. For instance, in the online Appendix I consider an example with two major powers, one more capable than the other. An alternative extension allows us to consider other types of interventions short of a full-on military campaign, such as providing support to insurgents and rebels, or conducting targeted bombing raids to destabilize a regime... these interventions fit within the theory by considering a lower launching cost of intervention than the large one (determined by g) in the current model. In a further extension, we may wish to study a richer strategic environment with multiple target countries, each of them potentially subject to foreign intervention.

The ultimate goal of the theory is to provide a tool that can help us predict and analyze future interventions, and the likely international reaction to them. For instance, in the years from 2017 to 2019, the US has threatened to intervene to topple the Assad regime in Syria, the Kim Jong-un regime in North Korea, and the Iranian regime. These and other conflict scenarios can be studied in this framework. Better calibration of the geographic matrix Λ , the preference profile, the expectations about the policy likely to be implemented after regime change, and the cost function of investments in conflict (which can vary by country) will allow decision makers to make better foreign policy decisions.

References

- [1] Aidt, Toke, Facundo Albornoz and Esther Hauk. “Foreign influence and domestic policy: a survey.” CESifo Working Paper Series 7567, CESifo Group Munich, 2019.
- [2] Belasco, Amy. “The Cost of Iraq, Afghanistan, and other global War on Terror operations since 9/11.” *Congressional Research Service* report, December 8, 2014.
- [3] Benson, Brett V., and Joshua D. Clinton. “Assessing the variation of formal military alliances.” *Journal of Conflict Resolution* 60.5: 866-898 (2016).
- [4] Benson, Brett V. *Constructing International Security: Alliances, Deterrence, and Moral Hazard*. Cambridge University Press, 2012.
- [5] Bloch, Francis, Santiago Sánchez-Pagés, and Raphaël Soubeyran. “When does universal peace prevail? Secession and group formation in conflict.” *Economics of Governance* 7(1): 3-29 (2006).
- [6] Bonfatti, Roberto. “An economic theory of foreign interventions and regime change.” *Canadian Journal of Economics* 50.1: 306-339 (2017).
- [7] Bueno De Mesquita, Bruce, James D. Morrow, Randolph M. Siverson, and Alastair Smith. “An institutional explanation of the democratic peace.” *American Political Science Review* 93(4): 791-807 (1999).
- [8] Esteban, Joan, and Debraj Ray. “Conflict and distribution.” *Journal of Economic Theory* 87(2): 379-415 (1999).
- [9] Fearon, James D. “Rationalist explanations for war.” *International Organization* 49.3: 379-414 (1995).
- [10] Gallop, Max. “More dangerous than dyads: how a third party enables rationalist explanations for war.” *Journal of Theoretical Politics* 29(3): 353-381 (2017).

- [11] Gratton, Gabriele and Bettina Klose. "Moderating conflicts with radical hardliners." Working paper, version April 20, 2017.
- [12] Jackson, Matthew O., and Massimo Morelli. "The reasons for wars: an updated survey." *The Handbook on the Political Economy of War* 34 (2011).
- [13] Kreps, Sarah E. *Coalitions of Convenience: United States Military Interventions after the Cold War*. Oxford University Press on Demand (2011).
- [14] Mas-Colell, Andreu; Michael D. Whinston and Jerry R. Green. *Microeconomic Theory*. Oxford U. Press (1995).
- [15] Niou, Emerson MS, and Peter C. Ordeshook. "Stability in anarchic international systems." *American Political Science Review* 84(4): 1207-1234 (1990).
- [16] Olson, Mancur, and Richard Zeckhauser. "An economic theory of alliances." *Review of Economics and Statistics* 48(3): 266-279 (1966).
- [17] Pompeo, Michael R. "Confronting Iran." *Foreign Affairs*. 97: 60-71 (2018).
- [18] Powell, Robert. "Bargaining theory and international conflict." *Annual Review of Political Science* 5: 1-30 (2002).
- [19] Powell, Robert. "Why some persistent problems persist." *American Political Science Review*, forthcoming.
- [20] Rubinstein, Ariel. "Perfect equilibrium in a bargaining model." *Econometrica* 50(1): 97-109 (1982).
- [21] Skaperdas, Stergios. "On the formation of alliances in conflict and contests." *Public Choice* 96(1-2): 25-42 (1998).
- [22] Wagner, R. Harrison. "The theory of games and the balance of power." *World Politics* 38(4): 546-576 (1986).

- [23] Wagner, R. Harrison. "Peace, war, and the balance of power." *American Political Science Review* 88(3): 593-607 (1994).
- [24] Walt, Stephen M. *The Origins of Alliances*. Ithaca: Cornell University Press (1987).
- [25] Waltz, Kenneth. *Theory of International Politics*. Reading: Addison-Wesley (1979).
- [26] Wolford, Scott. *The Politics of Military Coalitions*. Cambridge University Press (2015).
- [27] Yacoubian, Mona. "Critical junctures in United States policy toward Syria. An assessment of the counterfactuals." *Simon-Skjoldt Center for the Prevention of Genocide Series of Occasional Papers* 3 (2017).

Appendix

This is the Appendix to the paper “Regime Change.” This version is from August 2019.

An Example with a Major Power and a Lesser Power

I illustrate the model in action in an example with two major powers $\{1, 2\}$, and one rogue country 0, loosely inspired by the strategic interaction between the foreign policies of the US (a bigger power), the UK (a smaller power with more moderate policy preferences) and a target country. I assume countries have Euclidean preferences over a unidimensional policy space.

Example 1 Consider a game with $N = \{0, 1, 2\}$, $\theta_1 = \frac{1}{2}$, $\theta_2 = \theta_0 = \frac{1}{4}$; $p_{1,0} = 0$, $p_{2,0} = \hat{p}_2 = \frac{1}{4}$ and $\hat{p}_0 = 1$; $\lambda_{0,1} = \lambda_{0,2} = \frac{1}{2}$, and $v_{i,0}(x) = |x - \hat{p}_i|$ for each $i \in \{0, 1, 2\}$. Assume $g = \frac{1}{10}$ and $c : [g, 1) \rightarrow \mathbb{R}_+$ is given by $c(w) = -\ln\left(\frac{1-w}{1-g}\right) - \frac{(w-g)}{1-g}$ so $c'(w) = \frac{1}{1-w} - \frac{1}{1-g}$ and $c''(w) = \frac{1}{(1-w)^2} > 0$ for any $w \in [g, 1)$.²⁵

I solve this example numerically,²⁶ identifying the set of pure strategy equilibria as a function of the connectedness parameter $\phi \in (0, 1)$.

Consistent with the general result in Proposition 9, if countries are not susceptible to each others policies (ϕ very low), there are no interventions and peace holds; whereas, if they are very connected (ϕ very high) there is an intervention and conflict with certainty. For intermediate ranges of connectedness, we have multiple equilibria, as there is an equilibrium in which either major power launches a unique intervention, but it's not an equilibrium for both of them to launch. The two major countries could also fail to coordinate in one of these two equilibria, and end up in a mixed strategy equilibrium that is worse for both of them.

An alternative timing of the game resolves this -not very realistic- coordination failure. While both $\{1, 2\}$ are major powers capable of launching an intervention in the rogue nation, Country 1 is more powerful. Suppose (as in the case of the US vis a vis the UK), Country 1 has superior intervention technology, which allows Country 1 to intervene faster. This makes Country 1 able to have both first and last mover advantage: if an intervention must be launched by a fixed date $T > 0$, and it takes $t_i > 0$ units of time for Country i to launch an intervention, with $t_1 < t_2$, then Country 1 can launch at t_1 before Country 2 intervenes, and it can also launch at $T - t_1$, when it is too late for Country 2 to intervene. I simplify this timing to consider the following discrete version.

Sequential timing game. At the build up stage, Country 1 decides whether to intervene or to wait; Country 2 observes this decision and makes its own final decision to intervene or

²⁵Note that rescaling and translating the unit of measure by $\omega = \frac{w-g}{1-g}$ so that we measure not the total investment, but the fraction of the feasible budget $[0, 1 - g]$ that a country chooses to invest (in addition to the investment g), then the cost function can be equivalently expressed as $C : [0, 1) \rightarrow \mathbb{R}_+$ given by $C(w) = -\ln(1 - \omega) - \omega$.

²⁶A Mathematica .nb file containing all the necessary calculations is available from the author.

not; if Country 1 had initially waited, it observes Country 2's decision and makes its own final decision to intervene. Simultaneously, and without observing any of these decisions, the rogue chooses its defences d_0 . If at least one of the two major powers launches an intervention, a conflict stage ensues, in which each of the major powers chooses how much to support each ongoing intervention.

Sequential timing eliminates the coordination problem: if Country 1 prefers to be the one to launch an intervention, it gets to launch it; but if Country 1 prefers that Country 2 to be one launching it, then Country 2 gets to choose whether to launch it, or to push the game to a branch in which Country 1 is the only one that can act. So Country 1 is the one to launch if it prefers to be the one launching a unique intervention, and Country 2 is the one to launch if both countries prefer that Country 2 do so. If they both prefer to free-ride on the other, then Country 2 gets its wish, and Country 1 launches an intervention.

The results, as a function of the connectedness parameter $\phi \in [0, 1]$ are summarized in the following claim.

Claim 12 *In Example 1's game, with sequential timing, the equilibria are as follows.*

i) If $\phi \in [0, 0.016]$, the unique equilibrium is a peace equilibrium: there are no interventions and no regime change. The target country invests only the minimum necessary g to run its government.

ii) If $\phi \in [0.017, 0.050]$, there is no pure equilibrium. There is a mixed strategy equilibrium in which Country 1 intervenes with positive probability, and Country 2 does not intervene.

iii) If $\phi \in (0.050, 1]$, there is a unique pure equilibrium, in which Country 1 intervenes, and Country 2 does not intervene.

Table 2 notes the investments in defence by the rogue, the investments in support of Country 1's intervention by Country 1 and Country 2, and the probability of regime change, for some values of ϕ .

ϕ	Defence d_0	Offence (r_1, r_2)	Pr[Regime Change]
0.2	0.428	(0.090, 0.040)	0.34
0.4	0.423	(0.139, 0.056)	0.44
0.6	0.416	(0.172, 0.068)	0.50
1	0.406	(0.218, 0.086)	0.56

Table 2: Equilibria with an Intervention by Country 1.

Notice that Country 1 incurs the lion's share of the costs. Since it is the bigger country, its collective-action problem is small; Country 2, on the other hand, largely free-rides on the efforts of Country 1, contributing only a small amount.

Proofs

For any subset $I \subseteq M$ of major powers that launch an intervention, and for any $j \in N \setminus \{0\}$, let $b(j) \in I$ denote the intervention most preferred by Country j , that is, $b(j) = \arg \max_{i \in I} v_{j,0}(p_{i,0})$.

Lemma 13 *For any $\hat{d} \equiv (\hat{d}_0, \hat{d}_1, \hat{d}_2, \dots, \hat{d}_n) \in [g, 1] \times [0, 1 - g]^n$, consider the conflict subgame that follows after countries choose the vector of defence investments $d = \hat{d}$ at the build up stage and a non-empty subset of major powers $I \subseteq M$ each launches an independent intervention. Then any strategy r_j such that $r_{j,h} > 0$ for some $h \in K \setminus \{b(j)\}$ is strictly dominated, and, further, if $v_{j,0}(p_{0,b(j)}) \leq v_{j,0}(\hat{p}_0)$, then any strategy r_j such that $r_{j,b(j)} > 0$ is strictly dominated as well.*

Proof. Suppose that $r_{j,h} > 0$ for some $h \in K \setminus \{b(j)\}$. Then by deviating to $\tilde{r}_{j,b(j)} = r_{j,b(j)} + r_{j,h}$ and $\tilde{r}_{j,h} = 0$, and keeping all other investments constant, Country j strictly improves its expected policy outcome at no additional cost, and hence the deviation is strictly profitable; since this holds irrespective of the actions of other players, the original strategy is strictly dominated. Further, if $v_{j,0}(p_{0,b(j)}) \leq v_{j,0}(\hat{p}_0)$ and $r_{j,b(j)} > 0$, by deviating to $\tilde{r}_{j,b(j)} = 0$, Country j attains a weakly better expected policy outcome, at a strictly lower cost, again for any strategy profile played by other players, so the original strategy is strictly dominated. ■

Lemma 14 *For any $\hat{d} \equiv (\hat{d}_0, \hat{d}_1, \hat{d}_2, \dots, \hat{d}_n) \in [g, 1] \times [0, 1 - g]^n$, consider the restricted conflict subgame that follows after countries choose the vector of defence investments $d = \hat{d}$ at the build up stage and a subset $I \subseteq M$ of major powers each launches an independent intervention. Let w_i be the utility function of i in this restricted subgame. Then w_i is continuous in r , and is quasiconcave in r_i over the subdomain of strategies that are not strictly dominated for i .*

Proof. The cost function c is continuous in r . Notice that because $d_0 \geq g > 0$, the contest function 1 is also continuous in r . It follows that the utility function w_i is continuous in r .

By Lemma 13, for each $j \in N$, within the set of strategies that are not strictly dominated, Country j only supports her most preferred intervention $b(j)$, and only if it is better than the status quo. I show that w_j is strictly concave in r_j , for r_j within the set of strategies that are not strictly dominated.

For each $j \in N \setminus \{0\}$ such that $v_{j,0}(p_{0,b(j)}) > v_{j,0}(\hat{p}_0)$, the marginal benefit for Country j of investing in $r_{j,b(j)}$ is proportional to the marginal probability of affecting the outcome by

contributing more, which is equal to

$$\begin{aligned}
\frac{\partial}{\partial r_{j,b(j)}} \left(\frac{\theta \cdot r_{\cdot,b(j)}}{\sum_{i \in I} \theta \cdot r_{\cdot,i} + \theta \cdot \hat{d}} \right) &= \frac{\theta_j \left(\sum_{i \in I} \theta \cdot r_{\cdot,i} + \theta \cdot \hat{d} \right) - \theta r_{\cdot,b(j)} \theta_j}{\left(\sum_{i \in I} \theta \cdot r_{\cdot,i} + \theta \cdot \hat{d} \right)^2} \\
&= \theta_j \frac{\sum_{i \in I \setminus \{b(j)\}} \theta \cdot r_{\cdot,i} + \theta \cdot \hat{d}}{\left(\sum_{i \in I} \theta \cdot r_{\cdot,i} + \theta \cdot \hat{d} \right)^2}, \tag{10}
\end{aligned}$$

which is strictly decreasing:

$$\frac{\partial^2}{(\partial r_{j,b(j)})^2} \left(\frac{\theta \cdot r_{\cdot,b(j)}}{\sum_{i \in I} \theta \cdot r_{\cdot,i} + \theta \cdot \hat{d}} \right) = \frac{-2\theta_j^2 \left(\sum_{i \in I \setminus \{b(j)\}} \theta \cdot r_{\cdot,i} + \theta \cdot \hat{d} \right)}{\left(\sum_{i \in I} \theta \cdot r_{\cdot,i} + \theta \cdot \hat{d} \right)^3} < 0.$$

Hence the benefit of supporting the intervention is strictly concave, and hence it is strictly quasiconcave. The sum of the benefit plus $-c$ is then strictly quasiconcave as well. ■

Lemma 15 *Any conflict subgame of the game has a unique equilibrium, and this unique equilibrium is in pure strategies.*

Proof. We first note that it suffices to show that any restricted conflict subgame, in which players are restricted to not choose strictly dominated actions, has a unique equilibrium and this equilibrium is in pure strategies. If this restricted game has a unique equilibrium, lifting the restriction does not generate any additional equilibrium, since no equilibrium involves strictly dominated strategies. Thus, it suffices to consider this restricted game, and in this restricted game, each country's utility function is quasiconcave in investment in support of interventions (Lemma 14). Since the utility functions are continuous in the profile of investments r , and quasiconcave in own investment r_i , (Lemma 14), a pure strategy equilibrium of the conflict subgame exists (Glicksberg [3]). I show uniqueness by contradiction.

Assume that r^* and r^{**} are two equilibria of a conflict subgame and without loss of generality, assume that $\sum_{i \in I} \theta \cdot r_{\cdot,i}^{**} \geq \sum_{i \in I} \theta \cdot r_{\cdot,i}^*$. Assume (absurd), that there exists $j \in N$

such that $r_{j,b(j)}^{**} > r_{j,b(j)}^*$. Since $\sum_{i \in I} r_{j,i}^{**} = r_{j,b(j)}^{**}$ and $\sum_{i \in I} r_{j,i}^* = r_{j,b(j)}^*$ (by Lemma 13), $\sum_{i \in I} r_{j,i}^{**} >$

$\sum_{i \in I} r_{j,i}^*$ and hence the marginal cost for i of investing in support of $b(j)$'s intervention is

strictly greater given r^{**} than given r^* . Since r^{**} and r^* are equilibria and $r_{j,b(j)}^{**} > 0$, the marginal benefit of investing must be strictly greater as well, and hence, $r_{h',b(h)}^{**} \geq r_{h',b(h)}^*$ for any country j' such that $b(j') = b(j)$, and thus $\theta \cdot r_{\cdot,b(j)}^{**} > \theta \cdot r_{\cdot,b(j)}^*$. However, from Expression 10, the marginal benefit for Country j of investing in support of $b(j)$'s intervention is proportional to

$$\left(\frac{\sum_{i \in I \setminus \{b(j)\}} \theta \cdot r_{\cdot,i} + \theta \cdot \hat{d}}{\sum_{i \in I} \theta \cdot r_{\cdot,i} + \theta \cdot \hat{d}} \right) \left(\frac{1}{\sum_{i \in I} \theta \cdot r_{\cdot,i} + \theta \cdot \hat{d}} \right).$$

Since

$$\sum_{i \in I} \theta \cdot r_{\cdot,i}^{**} \geq \sum_{i \in I} \theta \cdot r_{\cdot,i}^* \implies \frac{1}{\sum_{i \in I} \theta \cdot r_{\cdot,i}^{**} + \theta \cdot \hat{d}} \leq \frac{1}{\sum_{i \in I} \theta \cdot r_{\cdot,i}^* + \theta \cdot \hat{d}},$$

it follows that in order for the marginal benefit for Country j of investing in support of $b(j)$'s intervention to be strictly greater under r^{**} than under r^* , it must be that

$$\left(\frac{\sum_{i \in I \setminus \{b(j)\}} \theta \cdot r_{\cdot,i}^{**} + \theta \cdot \hat{d}}{\sum_{i \in I} \theta \cdot r_{\cdot,i}^{**} + \theta \cdot \hat{d}} \right) > \left(\frac{\sum_{i \in I \setminus \{b(j)\}} \theta \cdot r_{\cdot,i}^* + \theta \cdot \hat{d}}{\sum_{i \in I} \theta \cdot r_{\cdot,i}^* + \theta \cdot \hat{d}} \right),$$

or equivalently,

$$\frac{\theta \cdot r_{\cdot,b(j)}^{**}}{\sum_{i \in I} \theta \cdot r_{\cdot,i}^{**} + \theta \cdot \hat{d}} < \frac{\theta \cdot r_{\cdot,b(j)}^*}{\sum_{i \in I} \theta \cdot r_{\cdot,i}^* + \theta \cdot \hat{d}},$$

that is, the share of all investments in support of interventions that accrues to the intervention by $b(j)$ must be lower under r^{**} than under r^* . Since for any $h \in I$ such that $\theta \cdot r_{\cdot,h}^{**} > \theta \cdot r_{\cdot,h}^*$, it must be $\frac{\theta \cdot r_{\cdot,h}^{**}}{\sum_{i \in I} \theta \cdot r_{\cdot,i}^{**} + \theta \cdot \hat{d}} < \frac{\theta \cdot r_{\cdot,h}^*}{\sum_{i \in I} \theta \cdot r_{\cdot,i}^* + \theta \cdot \hat{d}}$; it follows that for any $h' \in I$ such

that $\theta \cdot r_{\cdot,h'}^{**} \leq \theta \cdot r_{\cdot,h'}^*$, it must also be $\frac{\theta \cdot r_{\cdot,h'}^{**}}{\sum_{i \in I} \theta \cdot r_{\cdot,i}^{**} + \theta \cdot \hat{d}} < \frac{\theta \cdot r_{\cdot,h'}^*}{\sum_{i \in I} \theta \cdot r_{\cdot,i}^* + \theta \cdot \hat{d}}$, and thus aggregating,

$\sum_{i \in I} \theta \cdot r_{\cdot,i}^{**} < \sum_{i \in I} \theta \cdot r_{\cdot,i}^*$. We reach a contradiction: for any $h \in I$ such that $\theta \cdot r_{\cdot,h}^{**} > \theta \cdot r_{\cdot,h}^*$, $\sum_{i \in I} \theta \cdot r_{\cdot,i}^{**} < \sum_{i \in I} \theta \cdot r_{\cdot,i}^*$ implies $\frac{\theta \cdot r_{\cdot,h}^{**}}{\sum_{i \in I} \theta \cdot r_{\cdot,i}^{**} + \theta \cdot \hat{d}} > \frac{\theta \cdot r_{\cdot,h}^*}{\sum_{i \in I} \theta \cdot r_{\cdot,i}^* + \theta \cdot \hat{d}}$, a contradiction. Thus, there does

not exist $j \in N$ such that $r_{j,b(j)}^{**} > r_{j,b(j)}^*$. Hence, $r_{j,b(j)}^{**} \leq r_{j,b(j)}^*$ for each $j \in N$, which together with $r_{j,i}^* = r_{j,i}^{**} = 0$ for any $j \in N$ and for any $i \neq b(j)$ (by Lemma 13) and with $\sum_{i \in I} \theta \cdot r_{\cdot,i}^{**} \geq$

$\sum_{i \in I} \theta \cdot r_{\cdot,i}^*$, implies $r^* = r^{**}$ so the equilibrium of the conflict substage is unique. ■

Proof of Proposition 1. Any conflict subgame has a unique equilibrium, and this equilibrium is in pure strategies (Lemma 15). Consider the Nash equilibrium correspondence of the conflict stage subgame, as a function of the first stage actions. This Nash equilibrium correspondence of the conflict subgame is upper hemi-continuous with respect to changes in first-stage actions (Borgers [1], or see Fudenberg and Tirole section 1.3 for a textbook exposition, and particularly pages 31 and 32). Since the equilibrium of this conflict subgame is unique, the Nash equilibrium correspondence is a function, and thus it is continuous. Therefore, the expected utility of Country j is continuous in the actions in the build-up stage.

For any $j \in N$, Country j 's strategy set of j is a compact set. Since the cost function c is continuous, the contest function is continuous, and the solution to the second stage subgame is also continuous in the first stage actions, it follows that the utility function of the full game for each player is continuous. Therefore, an equilibrium of the whole game exists (Glicksberg [3]). ■

For any game with a hegemon, and for any $x \in \mathbb{R}_{++}$ such that $\theta \cdot d = x$, if the hegemon intervenes, since the hegemon's is the only possible intervention, we can drop without ambiguity the subindex labeling the intervention, and we can let $r(x)$ denote the column vector of investments $r_{\cdot,1}$ in support of the intervention. I use this notation throughout the proofs of the results in this section.

Proof of Lemma 2. In an equilibrium with peace, $a_1 = 0$, $d_0 = g$, and $d_j = 0$ for any $j \in N \setminus \{0\}$. If the hegemon deviates to $a_1 = 1$, then d_0 remains fixed at $d_0 = g$, and at the conflict stage of the game, for each $j \in N \setminus \{0\}$, Country j plays $r_j(g)$ and the probability that $p_0 = p_{1,0}$ is

$$\frac{\theta \cdot r(g)}{\theta \cdot r(g) + \theta_0 g}.$$

If inequality (3) holds, the hegemon prefers peace to this probabilistic outcome. So the equilibrium with peace holds. If, on the contrary, (3) does not hold, then the hegemon prefers to deviate.

To show uniqueness, consider any other strategy profile in which the hegemon intervenes with positive probability $\rho \in (0, 1]$. Let z be an optimal best response by the rogue to a probability ρ of intervention and in anticipation of investments $r_j(z)$ by each $j \in N \setminus \{0\}$. That is, z and $r_j(z)$ by each $j \in N \setminus \{0\}$ constitute a partial equilibrium subject to an intervention that occurs with probability ρ . Notice that in the special case with a hegemon, the contest function (1) simplifies to

$$\frac{\theta \cdot r_{\cdot,1}}{\theta \cdot r_{\cdot,1} + \theta \cdot d},$$

or equivalently, dropping the subindex for the hegemon (a more convenient notation that is unambiguous in this section, where the hegemon is the unique major power),

$$\frac{\theta \cdot r}{\theta \cdot r + \theta \cdot d} \tag{11}$$

Notice that $z > g$, and

$$\frac{\theta \cdot r(z)}{\theta \cdot r(z) + \theta_0 z} < \frac{\theta \cdot r(g)}{\theta \cdot r(g) + \theta_0 g}, \quad (12)$$

where this inequality holds because for each $j \in N \setminus \{0\}$, the derivative of the contest function (11) with respect to r_j strictly decreases if r and d are scaled up, and for a fixed d it also decreases in $\theta \cdot r$, so if inequality (12) does not hold and hence $r(z) \geq \frac{z}{g}r(g) > r(g)$, then for each $j \in N \setminus \{0\}$ such that $r_j(z) > r_j(g)$, Country j prefers to deviate to some $r_j \leq r_j(g)$, which implies that inequality (12) must hold.

Suppose $r_1(z) \geq r_1(g)$. Then subject to an intervention, the utility for the hegemon is strictly lower than subject to an intervention in the subgame in which $d_0 = g$ (the cost of supporting the intervention is higher, and the probability of success is strictly lower), which is strictly lower than the utility of no intervention. So the hegemon strictly prefers to not intervene.

Suppose $r_1(z) < r_1(g)$. This means that the marginal cost of supporting the intervention is strictly lower following $d_0 = z$ than following $d_0 = g$, which (by equilibrium conditions) implies that the marginal benefit from supporting the intervention is strictly lower, which means that the marginal probability that the intervention succeeds is lower (because all other factors are unchanged), which means that for each $j \in \{2, \dots, n\}$, the marginal benefit of supporting the intervention is lower, which implies that the marginal cost must also be strictly lower, which implies that $r_j(z) < r_j(g)$ for each $j \in \{2, \dots, n\}$, which implies that the incentives to launch an intervention are strictly worse for the hegemon given $d_0 = z$ than given $d_0 = g$ and thus the hegemon does not launch an intervention. ■

Lemma 16 *A pure subgame perfect equilibrium of game $\tilde{\Gamma}$ exists.*

Proof. For any $z \in [g, 1]$, a unique equilibrium exists in the conflict subgame of game $\tilde{\Gamma}$ that follows after the rogue invests $d_0 = z$, and this equilibrium is in pure strategies (Lemma 15). Under subgame perfection, game $\tilde{\Gamma}$ can then be solved as an individual optimization problem by the rogue over the choice of d_0 . Since the solution to any subgame that follows after the rogue chooses d_0 is continuous in d_0 , and the utility function of the rogue country is continuous, it follows that the utility that the rogue obtains as a function exclusively of d_0 under the assumption of subgame perfection, is as well continuous. Thus, it attains a maximum. This maximum, together with the investments in support of the intervention that follow it, constitute a subgame perfect equilibrium of game $\tilde{\Gamma}$. ■

Proof of Lemma 3. Step 1. Consider a strategy profile in which the hegemon intervenes, and all agents choose investment actions that are equilibrium actions in the restricted investment game $\tilde{\Gamma}$. Suppose inequality (4) does not hold. Then, the hegemon prefers to deviate from the strategy profile under consideration to not intervene, and thus the strategy profile is not an equilibrium of game Γ .

Step 2: Consider a strategy profile in which the hegemon intervenes, $d_0 = \tilde{d}_0$, and for each $j \in N \setminus \{0\}$, $d_j = 0$ and $r_j = r_j(\tilde{d}_0)$. By definition, $r_j = r_j(\tilde{d}_0)$ constitute mutual

best responses at the conflict subgame; $d_j = 0$ is optimal for any game $\Gamma \in \mathcal{G}$ because Country j prefers the intervention to succeed; $d_0 = \tilde{d}_0$ is by definition optimal for the rogue in expectation of an intervention, and of $d_j = 0$ and $r_j = r_j(\tilde{d}_0)$ for each $j \in N \setminus \{0\}$; and then given $d_j = 0$ and $r_j = r_j(\tilde{d}_0)$ for each $j \in N \setminus \{0\}$ and $d_0 = \tilde{d}_0$, if condition (4) holds, to intervene is optimal for the hegemon, and the intervention, together with these investments constitutes a pure subgame perfect equilibrium of game Γ . ■

Proof of Lemma 4. In order for the hegemon to be indifferent between intervening or not given that $d_0 = z$, it must be that

$$\phi \lambda_{0,1} |v_{1,0}(p_{1,0}) - v_{1,0}(\hat{p}_0)| \frac{\theta \cdot r(z)}{\theta \cdot r(z) + \theta_0 z} = c \left(g + \frac{\theta_0}{\theta_1} g + r_1(z) \right), \quad (13)$$

from which we obtain the equilibrium value of z in the mixed strategy equilibrium. In a mixed strategy equilibrium in which the hegemon intervenes with probability $\sigma \in (0, 1)$, the optimization problem of the rogue is

$$\max_{z \in [g, 1]} \sigma |v_{0,0}(p_{1,0})| \frac{\theta_0 z}{\theta \cdot r(z) + \theta_0 z} - c(z).$$

The first order condition is

$$\sigma |v_{0,0}(p_{1,0})| \frac{\theta_0 (\theta \cdot r(z) + \theta_0 z) - \theta_0 z \left(\sum_{j=1}^n \theta_j \frac{\partial}{\partial z} r_j(z) \right)}{(\theta \cdot r(z) + \theta_0 z)^2} = c'(z),$$

and this indifference condition pins down the probability of intervention σ .

If the peace condition (3) does not hold, then for $z = g$, the left hand side of (13) is strictly greater than the right hand side. As noted in Lemma 15, each conflict subgame has a unique pure equilibrium. Further, as noted in the existence proof of Proposition 1, the equilibrium of the conflict subgame as a function of first period actions is continuous. Thus, the left hand side of (13) is continuous, and the right hand side is continuous as well. If the intervention equilibrium condition (4) does not hold, then for $z = \tilde{d}_0$, the left-hand side of (13) is strictly smaller than the right hand side. Since the two sides are continuous for any $z \in [g, \tilde{d}_0]$, it follows by the Intermediate Value theorem that there is a value of z for which equality (13) holds. This is the value that sustains the mixed strategy equilibrium. ■

Proof of Proposition 5. Let (ω, ϕ) denote a list of parameter values, where ω is a list with a value for each parameter in the model aside from ϕ . That is, ω includes the vector of sizes θ , the list $v = (v_{0,0}, \dots, v_{n,0})$ of utility functions over policy in the target country, the proximity matrix Λ , the policy $p_{1,0}$ that would result in the target after regime change, the government size g and the cost function c . Let $\Omega \times (0, 1)$ denote the set of all possible parameter values. For any $(\omega, \phi) \in \Omega \times (0, 1)$, let $\Gamma^{\omega, \phi}$ denote the game with parameter values (ω, ϕ) . Fix any $\omega \in \Omega$ such that $\Gamma^{\omega, \phi} \in \mathcal{G}$, so the game features a hegemon and a rogue. Extend the model to allow for $\phi = 1$ and define $\phi_{Peace} \equiv \sup\{\phi \in (0, 1] : \Gamma^{\omega, \phi} \text{ has}$

a peace equilibrium} as the supremum of all values of ϕ for which game $\Gamma^{\omega,\phi}$ has a peace equilibrium. Note that because $(0, 1]$ is bounded, the supremum of $\{\phi \in (0, 1] : \Gamma^{\omega,\phi} \text{ has a peace equilibrium}\}$ exists, so ϕ_{Peace} is well-defined.

Part (i). Because the equilibrium correspondence is upper hemi continuous (see the proof of Proposition 1), game $\Gamma^{\omega,\phi_{Peace}}$ itself has a peace equilibrium. We want to show that game $\Gamma^{\omega,\phi}$ also has a peace equilibrium for any $\phi \in (0, \phi_{Peace}]$. That is (Lemma 2), we want to show that

$$\phi \lambda_{0,1}(v_{1,0}(p_{1,0}) - v_{1,0}(\hat{p}_0)) \frac{\theta \cdot r^\phi(g)}{\theta \cdot r^\phi(g) + \theta_0 g} \leq c \left(g + \frac{\theta_0}{\theta_1} g + r_1^\phi(g) \right)$$

for any $\phi \in (0, \phi_{Peace}]$, where the superscript denotes the game under consideration. Because $\Gamma^{\omega,\phi_{Peace}}$ has a peace equilibrium,

$$\phi_{Peace} \lambda_{0,1}(v_{1,0}(p_{1,0}) - v_{1,0}(\hat{p}_0)) \frac{\theta \cdot r^{\phi_{Peace}}(g)}{\theta \cdot r^{\phi_{Peace}}(g) + \theta_0 g} \leq c \left(g + \frac{\theta_0}{\theta_1} g + r_1^{\phi_{Peace}}(g) \right)$$

Assume that there exists $j \in N \setminus \{0\}$ such that in the conflict subgame, $r_j^\phi(g) > r_j^{\phi_{Peace}}(g)$. This implies that the marginal cost for Country j is strictly greater in the equilibrium of game $\Gamma^{\omega,\phi}$ than in the equilibrium of game $\Gamma^{\omega,\phi_{Peace}}$. Because Country j is optimizing, and chooses r_j^ϕ so that the marginal benefit of investment equals the marginal cost, this implies that the marginal benefit of investing for Country j is strictly greater in the equilibrium of game $\Gamma^{\omega,\phi}$ than in the equilibrium of game $\Gamma^{\omega,\phi_{Peace}}$. Formally,

$$\phi \lambda_{0,j}(v_{j,0}(p_{1,0}) - v_{j,0}(\hat{p}_0)) \frac{\theta_0 g}{(\theta \cdot r^\phi(g) + \theta_0 g)^2} > \phi_{Peace} \lambda_{0,j}(v_{j,0}(p_{1,0}) - v_{j,0}(\hat{p}_0)) \frac{\theta_0 g}{(\theta \cdot r^{\phi_{Peace}}(g) + \theta_0 g)^2}$$

or equivalently,

$$\frac{\phi}{\phi_{Peace}} > \frac{(\theta \cdot r^\phi(g) + \theta_0 g)^2}{(\theta \cdot r^{\phi_{Peace}}(g) + \theta_0 g)^2}, \quad (14)$$

which, since $\phi < \phi_{Peace}$, implies that

$$\theta \cdot r^\phi(g) < \theta \cdot r^{\phi_{Peace}}(g),$$

and it also implies that for any $h \in N \setminus \{0, j\}$,

$$\phi \lambda_{0,h}(v_{h,0}(p_{1,0}) - v_{h,0}(\hat{p}_0)) \frac{\theta_0 g}{(\theta \cdot r^\phi(g) + \theta_0 g)^2} > \phi_{Peace} \lambda_{0,h}(v_{h,0}(p_{1,0}) - v_{h,0}(\hat{p}_0)) \frac{\theta_0 g}{(\theta \cdot r^{\phi_{Peace}}(g) + \theta_0 g)^2},$$

that is, the marginal benefit of investing for Country h is strictly greater in the equilibrium of game $\Gamma^{\omega,\phi}$ than in the equilibrium of game $\Gamma^{\omega,\phi_{Peace}}$. Therefore, $r_1^\phi \geq r_1^{\phi_{Peace}}$, and if $r_1^{\phi_{Peace}} > 0$ then $r_1^\phi > r_1^{\phi_{Peace}}$, and further, $r_h^\phi > r_h^{\phi_{Peace}}$ for any $h \in \{2, \dots, n\}$.

For any $x \in [0, 1]^{n-1}$, let $BR_1^\phi(x)$ denote the best response of the hegemon to $(r_2, \dots, r_n) = x$ in the conflict subgame with parameter ϕ given $d_0 = g$. Note that $r_j^\phi > r_j^{\phi_{Peace}}$ for any

$j \in \{2, \dots, n\}$ implies $BR_1^\phi(r^\phi) \leq BR_1^\phi(r^{\phi_{Peace}})$ and $\phi < \phi_{Peace}$ implies $BR_1^\phi(x) \leq BR_1^{\phi_{Peace}}(x)$ for any $x \in [0, 1]^{n-1}$, so $r_1^\phi \equiv BR_1^\phi(r^\phi) \leq BR_1^\phi(r^{\phi_{Peace}}) \leq BR_1^{\phi_{Peace}}(r^{\phi_{Peace}}) = r_1^{\phi_{Peace}}$ so if $r_1^{\phi_{Peace}} > 0$ we obtain a contradiction. So assume that $r_1^\phi = r_1^{\phi_{Peace}} = 0$. Then $r_j^\phi(g) > r_j^{\phi_{Peace}}(g)$ for all $j \in \{2, \dots, n\}$ cannot be optimal for Country h for any $h \in \{2, \dots, n\}$: the marginal benefit of own investment r_h is strictly decreasing in total investment $\theta \cdot r$, so the marginal investment for Country h at investment level $r_h^\phi(g) > r_h^{\phi_{Peace}}(g)$ is strictly negative given parameter ϕ_{Peace} and given that for any $j \in N \setminus \{0, 1, h\}$ Country j invests $r_j^{\phi_{Peace}}(g)$, and thus it is also negative (and of greater absolute value) given ϕ_{Peace} and given that for any $j \in N \setminus \{0, 1, h\}$ Country j invests $r_j^\phi(g) > r_j^{\phi_{Peace}}(g)$, and therefore it is also negative (and of even greater absolute value) given $\phi < \phi_{Peace}$ and given that for any $j \in N \setminus \{0, 1, h\}$ Country j invests $r_j^\phi(g) > r_j^{\phi_{Peace}}(g)$. So it cannot be $r_h^\phi(g) > r_h^{\phi_{Peace}}(g)$ as established above, and therefore it must be that $r_h^\phi(g) \leq r_h^{\phi_{Peace}}(g)$ for any $h \in \{2, \dots, n\}$.

Then the incentives for the hegemon to launch an intervention are unambiguously worse in game $\Gamma^{\omega, \phi}$ than in game $\Gamma^{\omega, \phi_{Peace}}$: other countries would support the intervention less, so to attain any given probability of success would be costlier for the hegemon, and the utility benefit of succeeding would be strictly smaller because $\phi < \phi_{Peace}$ so that any expected policy gain would have a lesser positive utility effect.

Thus, if an intervention was not beneficial in game $\Gamma^{\omega, \phi_{Peace}}$, it is not beneficial in game $\Gamma^{\omega, \phi}$.

Part (ii) is by definition of ϕ_{Peace} : since no peace equilibrium exists for any $\phi > \phi_{Peace}$, and an equilibrium exists (Proposition 1) for any $\phi \in (0, 1)$, it follows that for any $\phi > \phi_{Peace}$, the equilibrium features a strictly positive probability of intervention. ■

Proof of Proposition 6. I first prove part (iii). Let (ω, ρ) denote a list of parameter values, where ω is a list with a value for each parameter in the model aside from ρ . That is, ω in this proof now includes the vector of sizes θ , the connectivity parameter ϕ , the proximity matrix Λ , the vector of ideal policies $(\hat{p}_1, \dots, \hat{p}_n)$ the policy $p_{1,0}$ that would result in the target after regime change, the government size g and the cost function c . Let $\Omega \times \mathbb{R}_+$ denote the set of all possible parameter values. For any $(\omega, \rho) \in \Omega \times \mathbb{R}_+$, let $\Gamma^{\omega, \rho}$ denote the game with parameter values (ω, ρ) . Fix any $\omega \in \Omega$ such that $\Gamma^{\omega, \rho} \in \mathcal{G}$, so the game features a hegemon and a rogue.

Note that since $\theta \cdot d$ is bounded, for any $r_1 \in (0, 1 - g - \frac{\theta_0}{\theta_1}g)$, there exists $\varepsilon(r_1) \in \mathbb{R}_+$ such that by intervening and investing r_1 , the hegemon can guarantee a probability at least $\varepsilon(r_1)$ that the intervention succeeds. The cost of intervening and investing r_1 in support of the intervention is bounded; whereas, as $\rho \rightarrow \infty$, the benefit of intervening successfully becomes arbitrarily large, and hence there is $\rho_{Conflict} \in \mathbb{R}_+$ such that the hegemon intervenes with certainty if $\rho > \rho_{Conflict}$. So part (iii) is established.

Next I prove part (i). Define $\phi_{Peace} \equiv \sup\{\rho \in [0, \rho_{Conflict}] : \Gamma^{\omega, \rho} \text{ has a peace equilibrium}\}$ as the supremum (over $[0, \rho_{Conflict}]$) of all values of ρ for which game $\Gamma^{\omega, \rho}$ has a peace equilibrium. Note that because $[0, \rho_{Conflict}]$ is bounded, the supremum exists, so ϕ_{Peace} is well-defined. Because the equilibrium correspondence is upper hemi continuous (see the proof

of Proposition 1), game $\Gamma^{\theta, \rho_{Peace}}$ itself has a peace equilibrium. We want to show that game $\Gamma^{\theta, \rho}$ also has a peace equilibrium for any $\rho \in (0, \rho_{Peace}]$. That is (Lemma 2), we want to show that

$$\begin{aligned}\phi\lambda_{0,1}(v_{1,0}(p_{1,0}) - v_{1,0}(\hat{p}_0)) \frac{\theta \cdot r^\rho(g)}{\theta \cdot r^\rho(g) + \theta_0 g} &\leq c \left(g + \frac{\theta_0}{\theta_1} g + r_1^\rho(g) \right) \\ \phi\lambda_{0,1}(-\mu + (2\hat{p}_n - \hat{p}_1 - \mu + \rho - \hat{p}_1)) \frac{\theta \cdot r^\rho(g)}{\theta \cdot r^\rho(g) + \theta_0 g} &\leq c \left(g + \frac{\theta_0}{\theta_1} g + r_1^\rho(g) \right) \\ \phi\lambda_{0,1}(2\hat{p}_n - 2\hat{p}_1 + \rho) \frac{\theta \cdot r^\rho(g)}{\theta \cdot r^\rho(g) + \theta_0 g} &\leq c \left(g + \frac{\theta_0}{\theta_1} g + r_1^\rho(g) \right)\end{aligned}$$

for any $\rho \in (0, \rho_{Peace}]$, where superscripts denote the game under consideration. Because $\Gamma^{\theta, \rho_{Peace}}$ has a peace equilibrium,

$$\phi\lambda_{0,1}(2\hat{p}_n - 2\hat{p}_1 + \rho) \frac{\theta \cdot r^{\rho_{Peace}}(g)}{\theta \cdot r^{\rho_{Peace}}(g) + \theta_0 g} \leq c \left(g + \frac{\theta_0}{\theta_1} g + r_1^{\rho_{Peace}}(g) \right)$$

Suppose $\sum_{j=2}^n \theta_j r_j^\rho(g) < \sum_{j=2}^n \theta_j r_j^{\rho_{Peace}}(g)$. Then the hegemon has a lessened incentive to launch an intervention, so non-intervention given ρ_{Peace} implies non-intervention given ρ , as desired.

Suppose $\sum_{j=1}^n \theta_j r_j^\rho(g) \geq \sum_{j=1}^n \theta_j r_j^{\rho_{Peace}}(g)$. Then the marginal incentive to contribute to support the intervention is strictly lower for every country given ρ than given ρ_{Peace} , so $r_1^\rho(g) \leq r_1^{\rho_{Peace}}(g)$ and $r_j^\rho(g) < r_j^{\rho_{Peace}}(g)$ for any $j \in \{2, \dots, n\}$, and therefore $\sum_{j=1}^n \theta_j r_j^\rho(g) < \sum_{j=1}^n \theta_j r_j^{\rho_{Peace}}(g)$, a contradiction.

So suppose $\sum_{j=1}^n \theta_j r_j^\rho(g) < \sum_{j=1}^n \theta_j r_j^{\rho_{Peace}}(g)$ and $\sum_{j=2}^n \theta_j r_j^\rho(g) \geq \sum_{i=2}^n \theta_i r_i^{\rho_{Peace}}(g)$, which implies $r_1^\rho(g) < r_1^{\rho_{Peace}}(g)$. Because the hegemon chooses r_1^ρ so that the marginal benefit of investment equals the marginal cost, this implies that the marginal benefit of investing is strictly lower in the equilibrium of game $\Gamma^{\theta, \rho}$ than in the equilibrium of game $\Gamma^{\theta, \rho_{Peace}}$. Formally,

$$\phi\lambda_{0,1}(2\hat{p}_n - 2\hat{p}_1 + \rho) \frac{\theta_1 \theta_0 g}{(\theta \cdot r^\rho(g) + \theta_0 g)^2} < \phi\lambda_{0,1}(2\hat{p}_n - 2\hat{p}_1 + \rho_{Peace}) \frac{\theta_1 \theta_0 g}{(\theta \cdot r^{\rho_{Peace}}(g) + \theta_0 g)^2}$$

or equivalently,

$$\frac{2\hat{p}_n - 2\hat{p}_1 + \rho}{2\hat{p}_n - 2\hat{p}_1 + \rho_{Peace}} < \frac{(\theta \cdot r^\rho(g) + \theta_0 g)^2}{(\theta \cdot r^{\rho_{Peace}}(g) + \theta_0 g)^2},$$

which, since $\hat{p}_j \geq \hat{p}_1$ for any $j \in N \setminus \{n, 1\}$, implies

$$\frac{2\hat{p}_n - 2\hat{p}_j + \rho}{2\hat{p}_n - 2\hat{p}_j + \rho_{Peace}} < \frac{(\theta \cdot r^\rho(g) + \theta_0 g)^2}{(\theta \cdot r^{\rho_{Peace}}(g) + \theta_0 g)^2}$$

and

$$\phi\lambda_{0,j}(2\hat{p}_n - 2\hat{p}_j + \rho) \frac{\theta_j\theta_0g}{(\theta \cdot r^\rho(g) + \theta_0g)^2} < \phi\lambda_{0,j}(2\hat{p}_n - 2\hat{p}_j + \rho_{Peace}) \frac{\theta_j\theta_0g}{(\theta \cdot r^{\rho_{Peace}}(g) + \theta_0g)^2};$$

that is, the marginal benefit of investing for Country j is strictly lower in the equilibrium of game $\Gamma^{\theta,\rho}$ than in the equilibrium of game $\Gamma^{\theta,\rho_{Peace}}$, and thus, the marginal cost must be lower as well, which implies $r_j^\rho(g) < r_j^{\rho_{Peace}}(g)$, and since this holds for any $j \in \{2, \dots, n\}$, it follows that $\sum_{i=2}^n \theta_i r_i^\rho(g) < \sum_{i=2}^n \theta_i r_i^{\rho_{Peace}}(g)$, a contradiction. So it must be that $\sum_{i=2}^n \theta_i r_i^\rho(g) < \sum_{i=2}^n \theta_i r_i^{\rho_{Peace}}(g)$. Then, as noted above, the hegemon has a lessened incentive to launch an intervention, so non-intervention given ρ_{Peace} implies non-intervention given ρ , as desired.

Part (ii) is by definition of ρ_{Peace} . ■

Proof of Proposition 7. Part (i). Let (ω, θ_0) denote a list of parameter values, where ω is a list with a value for each parameter in the model aside from θ_0 . That is, ω in this proof now includes the connectivity parameter ϕ , the proximity matrix Λ , the list of utility functions over policy $\{v_{j,0}\}_{j=0}^\infty$, the policy $p_{1,0}$ that would result in the target after regime change, the government size g , the cost function c , and the relative size $\hat{\theta}_j \equiv \frac{\theta_j}{1-\theta_j}$ for each country $j \in N \setminus \{0\}$. Let $\Omega \times (0, 1)$ denote the set of all possible parameter values. For any $(\omega, \theta_0) \in \Omega \times \mathbb{R}_{++}$, let $\Gamma^{\omega, \theta_0}$ denote the game with parameter values (ω, θ_0) . Fix any $\omega \in \Omega$ such that $\Gamma^{\omega, \rho} \in \mathcal{G}$, so the game features a hegemon and a rogue, and let \mathcal{G}^ω be the class of games with parameters ω , indexed by θ_0 . We run comparative statics on θ_0 within the class of games \mathcal{G}^ω .

Note that if $\theta_0 = 0$, launching an intervention in the rogue country would be costless, and for any strictly positive investment in support of the intervention, the intervention succeeds. For any $\phi > 0$, any $\lambda_{0,1} > 0$ and any $v_{1,0}(\hat{p}_0) < 0$, it would therefore be strictly beneficial to launch an intervention if $\theta_0 = 0$. By continuity of the cost function and the contest success function, the hegemon also has a strict incentive to intervene for any θ_0 sufficiently close to zero. Furthermore, for any $\pi \in (0, 1)$, as θ_0 converges toward zero, the cost of launching and supporting an intervention that would succeed with probability at least π converges to zero. Since the benefit of regime change is fixed in θ_0 , for any sufficiently low θ_0 , it is optimal to launch an intervention that would succeed with probability at least π . So part (i) is established.

Part (iii). Define $\theta_0^{Peace} \equiv \inf\{\theta_0 \in (0, 1) : \Gamma^{\omega, \theta_0} \text{ has a peace equilibrium}\}$ as the infimum of all values of θ_0 for which game $\Gamma^{\omega, \theta_0}$ has a peace equilibrium. Because the equilibrium correspondence is upper hemicontinuous, game $\Gamma^{\omega, \theta_0^{Peace}}$ itself has a peace equilibrium. We want to show that game $\Gamma^{\omega, \theta_0}$ also has a peace equilibrium for any $\theta_0 \in (\theta_0^{Peace}, 1)$. That is

(Lemma 2), we want to show that

$$\phi\lambda_{0,1}|v_{1,0}(\hat{p}_0) - v_{1,0}(p_{1,0})| \frac{\sum_{j=1}^n \hat{\theta}_j(1 - \theta_0)r_j^{\theta_0}(g)}{\sum_{j=1}^n \hat{\theta}_j(1 - \theta_0)r_j^{\theta_0}(g) + \theta_0 g} \leq c \left(g + \frac{\theta_0}{\hat{\theta}_1(1 - \theta_0)}g + r_1^{\theta_0}(g) \right)$$

for any $\theta_0 \in (\theta_0^{Peace}, 1]$, where superscripts denote the game under consideration. Because $\Gamma^{\omega, \theta_0^{Peace}}$ has a peace equilibrium,

$$\phi\lambda_{0,1}|v_{1,0}(\hat{p}_0) - v_{1,0}(p_{1,0})| \frac{\sum_{j=1}^n \hat{\theta}_j(1 - \theta_0^{Peace})r_j^{\theta_0^{Peace}}(g)}{\sum_{j=1}^n \hat{\theta}_j(1 - \theta_0^{Peace})r_j^{\theta_0^{Peace}}(g) + \theta_0^{Peace}g} \leq c \left(g + \frac{\theta_0^{Peace}}{\hat{\theta}_1(1 - \theta_0^{Peace})}g + r_1^{\theta_0^{Peace}}(g) \right)$$

Assume that given θ_0^{Peace} , an (off equilibrium) intervention would succeed with probability less than one half. Then an increase in the size of the rogue to $\theta_0 > \theta_0^{Peace}$ reduces the incentive of any other country to support the intervention, so they all reduce their investment. Hence the incentives to launch an intervention for the hegemon are unambiguously worse with θ_0 than with θ_0^{Peace} (formally, $\sum_{j=2}^n \hat{\theta}_j(1 - \theta_0)r_j^{\theta_0}(g)$ in support of the intervention is lower than $\sum_{j=2}^n \hat{\theta}_j(1 - \theta_0^{Peace})r_j^{\theta_0^{Peace}}(g)$ and $\theta_0 g$ against it is higher, so for any investment r_1 the probability of succeeding is lower, and furthermore the cost $\frac{\theta_0}{\hat{\theta}_1(1 - \theta_0)}g$ of launching an intervention is also higher). So if peace holds for parameter θ_0^{Peace} , it also holds for any $\theta_0 > \theta_0^{Peace}$.

Assume instead that given θ_0^{Peace} , an (off equilibrium) intervention succeeds with probability more than one half. A marginal increase in θ_0 reduces the probability that the intervention succeeds, incentivizing all countries that support the intervention to support it more. A counteracting effect, the decrease in the size θ_j for each $j \in N \setminus \{0\}$ as θ_0 increases, incentivizes all these countries to invest less. If the negative effect dominates, the argument in the previous paragraph applies: to intervene is unambiguously less attractive for the hegemon as θ_0 increases. If the positive effect dominates, $r_1^{\theta_0}(g) > r_1^{\theta_0^{Peace}}(g)$ and yet the probability that the intervention succeeds is lower in game Γ^{θ_0} than in game $\Gamma^{\theta_0^{Peace}}$. Thus, the benefit of intervention is lower, and the cost is higher: the intervention is strictly less attractive to the hegemon in game Γ^{θ_0} than in game $\Gamma^{\theta_0^{Peace}}$.

Part (ii) is by definition of θ_0^{Peace} . ■

Proof of Lemma 8. If condition 5 is not satisfied, then there exists $i \in M$ who strictly prefers to deviate from the peace equilibrium. If condition 5 is satisfied, no major power i wishes to deviate from the peace equilibrium.

If the target country is rogue, it means that for any $i \in M$, and for any $\hat{M} \subset M \setminus \{i\}$, major power i prefers any outcome in which at least one $j \in \hat{M}$ intervenes with positive

probability, over the peace equilibrium. So if major power i has no incentives to intervene given a peace equilibrium, then i also has no incentive to intervene given that at least one $j \in \hat{M}$ intervenes with positive probability. So there cannot be an equilibrium in which at least two major powers intervene with positive probability. An equilibrium in which only one major power intervenes with positive probability is ruled out by the proof of Lemma 2 with a hegemon. ■

Proof of Proposition 9. By Proposition 5, for each $i \in M$ there exists $\phi_{Peace}(i)$ such that if i were the hegemon, a peace equilibrium exists if and only if $\phi \leq \phi_{Peace}(i)$. Define $\phi_{Peace}^M = \min_{i \in M} \phi_{Peace}(i)$. Then, if $\phi \leq \phi_{Peace}^M$, no major power has an incentive to deviate from a peace equilibrium, so a peace equilibrium exists. Further, for each $i \in M$, define $w_i \in \mathbb{R}$ by $w_i \equiv \min\{v_{i,0}(\hat{p}_0), \{v_{i,0}(p_{0,j})\}_{j \in M}\}$. That is, w_i is the worst possible policy outcome for i . If

$$\phi \leq \frac{c \left(g + \frac{\theta_0}{\theta_i} g \right)}{\lambda_{0,i} |w_i|},$$

then to not intervene is a dominant strategy for major power j . Hence, if $\phi \leq \min_{i \in M} \left\{ \frac{c \left(g + \frac{\theta_0}{\theta_i} g \right)}{\lambda_{0,i} |w_i|} \right\} \equiv \phi_{Unique}^M$, then the unique equilibrium is a peace equilibrium.

If $\phi > \phi_{Peace}^M = \min_{i \in M} \phi_{Peace}(i)$, at least one major power would deviate from a peace equilibrium, so this equilibrium doesn't hold, but since an equilibrium exists (Proposition 1), in equilibrium an intervention occurs with strictly positive probability. ■

Proof of Proposition 10. Since the alliance optimizes by maximizing the aggregate welfare of its members, it operates as a single agent that at the conflict stage equalizes marginal cost of further investment to marginal benefit, and at the build up stage it launches an intervention if and only if doing so generates a greater expected sum of utilities to its members than not doing so. Define an agent a with size $\sum_{j \in A} \theta_j$ and policy preferences $v_{a,0}(p_0) = \sum_{i \in A} \frac{\theta_i}{\sum_{j \in A} \theta_k} v_{i,0}(p_0)$. Define the game with countries $(N \setminus A) \cup \{a\}$, and let a be the hegemon with an ability to intervene. Apply the proof of Proposition 5 to this game. ■

Proof of Proposition 11. Given any $d_0 \in [g, 1]$, assume an intervention occurs. If there is no alliance, for any $j \in N \setminus \{0, 1\}$, the hegemon and Country j respectively solve

$$\begin{aligned} & \max_{r_1 \in [0, 1-g-\frac{\theta_0}{\theta_1}g]} \phi \lambda_{0,1} (v_{1,0}(p_{1,0}) - v_{1,0}(\hat{p}_0)) \frac{\theta \cdot r}{\theta \cdot r + \theta_0 d_0} - c \left(g + \frac{\theta_0}{\theta_1} g + r_1 \right) \\ & \max_{r_j \in [0, 1-g]} \phi \lambda_{0,j} (v_{j,0}(p_{1,0}) - v_{j,0}(\hat{p}_0)) \frac{\theta \cdot r}{\theta \cdot r + \theta_0 d_0} - c(g + r_j) \end{aligned}$$

with first order conditions

$$\begin{aligned}\phi\lambda_{0,1}(v_{1,0}(p_{1,0}) - v_{1,0}(\hat{p}_0))\frac{\theta_1\theta_0d_0}{(\theta \cdot r + \theta_0d_0)^2} &= c' \left(g + \frac{\theta_0}{\theta_1}g + r_1 \right) \\ \phi\lambda_{0,j}(v_{j,0}(p_{1,0}) - v_{j,0}(\hat{p}_0))\frac{\theta_j\theta_0d_0}{(\theta \cdot r + \theta_0d_0)^2} &= c' (g + r_j)\end{aligned}\quad (15)$$

whereas the alliance solves

$$\max_{\{r_i^A\}_{i \in A}} \left\{ \phi \sum_{i \in A} \left(\frac{\theta_i \lambda_{0,i}(v_{i,0}(p_{1,0}) - v_{i,0}(\hat{p}_0))}{\sum_{k \in A} \theta_k} \right) \frac{\theta \cdot r^A}{\theta \cdot r^A + \theta_0 d_0} - \frac{\theta_1 c \left(g + \frac{\theta_0}{\theta_1} g + r_1^A \right)}{\sum_{k \in A} \theta_k} - \sum_{i \in A \setminus \{1\}} \frac{\theta_i c (g + r_j^A)}{\sum_{k \in A} \theta_k} \right\},$$

where r^A denotes the vector of all investments in support of the intervention by all countries (not just alliance members) in the conflict subgame of the game with an alliance. The first order conditions are

$$\begin{aligned}\phi \sum_{i \in A} \theta_i \lambda_{0,i}(v_{i,0}(p_{1,0}) - v_{i,0}(\hat{p}_0))\frac{\theta_1\theta_0d_0}{(\theta \cdot r^A + \theta_0d_0)^2} &= \theta_1 c' \left(g + \frac{\theta_0}{\theta_1} g + r_1^A \right), \text{ and} \\ \phi \sum_{i \in A} \theta_i \lambda_{0,i}(v_{i,0}(p_{1,0}) - v_{i,0}(\hat{p}_0))\frac{\theta_j\theta_0d_0}{(\theta \cdot r^A + \theta_0d_0)^2} &= \theta_j c' (g + r_j^A) \quad \forall j \in A \setminus \{1\},\end{aligned}$$

if the solution is interior, and with $r_1 = 0$ at a corner solution. Equivalently, express the first order conditions as

$$\phi \sum_{i \in A} \theta_i \lambda_{0,i}(v_{i,0}(p_{1,0}) - v_{i,0}(\hat{p}_0))\frac{\theta_0d_0}{(\theta \cdot r^A + \theta_0d_0)^2} = c' \left(g + \frac{\theta_0}{\theta_1} g + r_1^A \right) = c' (g + r_i^A) \quad \forall j \in A \setminus \{1\}.$$

The left hand side of the alliance first order conditions for r_j^A is strictly greater than the corresponding left hand side of the first order conditions for r_j in the individual optimization problem of Country j (expression 15), and thus to preserve equality, the right hand solution must be greater as well, which, since c is strictly increasing, implies that $r_j^A > r_j$ for any $j \in A \setminus \{1\}$ and similarly $r_1^A \geq r_1$ with equality only at a corner solution with $r_1^A = 0$. That is, $r_j^A(d_0) > r_j(d_0)$ for any $j \in A \setminus \{1\}$ and $r_1^A(d_0) \geq r_1(d_0)$, for any $d_0 \in [g, 1]$. Given an increase in the alliance investment from $\sum_{i \in A} r_i$ to $\sum_{i \in A} r_i^A$, countries outside the alliance have a lessened incentive to invest in support of the intervention, and thus invest less (thus, part (a) of Proposition 11 is established). However, the incentive to invest is lessened only insofar as the total investment is greater, so it must be that $\sum_{i \in N \setminus \{0\}} r_i^A > \sum_{i \in N \setminus \{0\}} r_i$.

Thus, either the rogue invests more on its own defence, or else, subject to an intervention occurring, regime change is more likely (thus, part (b) is established).

Suppose $A = N \setminus \{0\}$. The alliance optimizes investments jointly, which by definition, must result in greater aggregate welfare for the alliance than choosing individual investments

separately. Suppose the alliance does not intervene in equilibrium. We want to show that then the hegemon does not intervene without an alliance. For any vector of investments $\{\tilde{r}_i\}_{i \in A}$ that do not correspond to the alliance's equilibrium investments in the conflict subgame, if the alliance does not intervene in equilibrium, then it would not intervene, subject to being constrained to choosing investments $\{\tilde{r}_i\}_{i \in A}$. In particular, it would not intervene subject to investing $\{r_{i,1}^\phi\}_{i \in A}$, which are the equilibrium values for the non-alliance game Γ^ϕ . For any $i \in A \setminus \{1\}$, $v_{i,0}(p_{1,0}) > v_{i,0}(\hat{p}_0)$ by assumption, and thus an intervention with investments $\{r_k^\phi\}_{k \in A \setminus \{i\}}$ and $r_i = 0$ is strictly beneficial for Country i . By optimality of i 's choices, it follows that an intervention with $\{r_k^\phi\}_{k \in A}$ is also strictly beneficial for Country i . Since the alliance would not intervene subject to investing $\{r_k^\phi\}_{k \in A}$, it follows that the aggregate welfare of this intervention for all countries in A is negative, and since it is positive for any $i \in A \setminus \{1\}$, it must be that the intervention is detrimental to the hegemon, so it would not launch it. So if the alliance would not intervene, the hegemon would not intervene (strictly, without indifference) without an alliance. Thus, $\gamma_{Peace}^A < \gamma_{Peace}$. ■

References

- [1] Börgers, Tilman. "Upper hemicontinuity of the correspondence of subgame-perfect equilibrium outcomes." *Journal of Mathematical Economics* 20: 89-106 (1991).
- [2] Fudenberg, Drew, and Jean Tirole. *Game Theory*. MIT Press. (1991).
- [3] Glicksberg, Irving L. "A further generalization of the Kakutani fixed point theorem, with application to Nash equilibrium points." *Proceedings of the American Mathematical Society* 3.1: 170-174 (1952).