Collective Responses to Rogue State Actions$^1$

Sanford C. Gordon
Department of Politics
New York University
sanford.gordon@nyu.edu

Catherine Hafer
Department of Politics
New York University
catherine.hafer@nyu.edu

Dimitri Landa
Department of Politics
New York University
dimitri.landa@nyu.edu

November 1, 2012

$^1$Very preliminary and incomplete. Please do not cite or circulate.
1 Introduction

International responses to provocative actions by “rogue states” take place in an environment characterized by two key features. The first is incomplete information about the ultimate preferences of other states: just how much they dislike the rogue state’s actions, and so, how far they would ultimately be willing to go in incurring costs to stop it or slow it down. The second is coordination by subsets of states in response to a rogue state’s actions – either within the framework of the United Nations or outside it, through “coalitions of the willing.” The focus of the present paper is on the effectiveness of such coordinated responses in an environment of incomplete information.

Our point of departure is the conjunction of two concerns about this environment raised by foreign policy makers. The first, most closely identified with John Bolton, the former U.S. Permanent Representative to the UN, is that the United Nations is an ineffective, and possibly even counterproductive, framework for international responses to rogue states. The reasons are two-fold: first, it leaves decision-making in the hands of less affected members – and perhaps even members who benefit from the rogue state’s actions. This effectively undermines the ability of adversely affected members – particularly the United States – to respond to the rogue states in a manner commensurate with its own interests. Second, the UN decision-making process reveals to rogue states information about affected states that can ultimately hurt the latter by empowering the former (Bolton 2012).

An alternative advocated by critics of the UN framework is one of “coalitions of the willing,” a term first used by Bill Clinton to describe international intervention in Bosnia in 1994, and which achieved greater prominence in the Bush administration’s run-up to the 2003 Iraq War. In this framework, adverse actions by rogue states legitimize “private” responses, and so adversely affected states can simply join a coalition that seeks to impose strong sanctions, perhaps militarily, on a rogue state.

This brings us to the second concern: that coalitions of the willing change the nature of the informational environment. In particular, the very possibility of such coalitions may alter the beliefs of the rogue states’ leaders in a way that undermines simplistic interpretations of what qualifies as “willing.” Put simply, the existence of a coalition of the willing exerts pressures on states to join it, quite apart from any side-deals between members. Such states may feel compelled to join
coalitions they preferred not be possible. Strictly voluntary coalitions of the willing, thus, may be comprised of the truly “willing” and those “compelled’ to join because of the fear of the signal that would be sent to the rogue state by a decision not to join. This intuition resonates with Tony Blair’s account of his decision to have the UK join the US-led coalition of the willing to take a military action against Saddam Hussein: once the US had made is clear that its intention to take the action, Blair became convinced that it was better that the UK join in, and, indeed, that was the argument he tried making to other European leaders.

We provide a model that captures these intuitions. In our model, states differ in their underlying costs associated with a rogue state’s behavior: while some states suffer strong negative disutilities, for other states, those disutilities may be lower. For still others, the rogue state’s actions may be beneficial. We consider states’ responses under three institutional scenarios: “pact,” in which states voluntarily pay the cost of participating in joint action (above and beyond the costs they impose on the rogue state privately); “strong UN,” under which any coordinated action by the international community requires UN authorization; and “weak UN with breakaway coalitions of the willing,” in which states may form a coalition to sanction the rogue state in the event that UN authorization is not granted. These three scenarios are contrasted with an “anarchic baseline,” under which no coordinated action among non-rogue states is possible.

We show that, from the ex ante perspective, both coalitions of the lead to a more adverse action by the rogue state than under the anarchic baseline; however, individual states may prefer one or the other to anarchy. A highly-motivated state that suspects others are not likely to be will strictly prefer a coalition of the willing to a UN environment. Other states that would prefer neither institution exist might nonetheless feel compelled to join either an enforcing coalition or a coalition of the willing. We show how these different institutions can give rise to different coalitions emerging to retaliate against provocative actions by rogue states, and how the expected actions of rogue states themselves can vary across institutions.
2 Background

2.1 Previous Research on Signaling and International Conflict

Our analysis lies at the intersection of three literatures in the field of international relations. The first concerns the extent to and conditions under which threats to escalate militarily can convey credible information about a state’s resolve during interstate crises. Fearon (1994) provides the seminal analysis, arguing that democracies are better able to signal resolve to adversaries than non-democracies because of the domestic political audience costs associated with capitulating. Schultz (1998) expands Fearon’s argument by incorporating electoral competition in democratic states explicitly: in that model, the presence of an opposition party can enhance the credibility of an incumbent’s messages to international adversaries: specifically, by lending support to the incumbent’s policy.\footnote{Smith (1998a) also considers the effect of the domestic political environment on the credibility of leaders’ foreign policy declarations.} Ramsay (2004) considers a model in which domestic opposition reveals credible information about an incumbent to rival states. And in Slantchev (2006), the media’s role in providing information to domestic constituents serves a secondary role in increasing the audience costs to leaders of failure.

The second literature concerns the use by states of the United Nations and, more specifically, the UN Security Council. In Voeten (2001), a credible outside option improves a state’s prospects for winning favorable votes in the Security Council, although uncertainty about the credibility of that option can lead to unilateral action on the path of play. Others have argued that the UN enhances the legitimacy of the use of force. Voeten (2005), for example, argues that an authorization by the UN Security Council indicates that the political fallout to a state from its use of force is likely to be limited. The mechanism that allows the UN to function in this role is a shared norm to “green light” states that cooperate with this norm and punish violators. Hurd (2008) argues that when states seek authorization from the UN Security Council to pursue their own objectives, they contribute to the Council’s authority.

The third body of work concerns the formation of alliances in the international system. Most relevant in the current context are models that consider the collective action and signaling issues inherent in the decision to form an alliance. Olson and Zeckhauser (1966) provide a collective action-
theoretic account of alliances, arguing that incentives to free-ride can undermine international alliances such as NATO. In the model of Morrow (1994), a state can use alliance formation as a costly signal of the disutility of fighting. States must trade off between the peacetime costs of alliance formation and the enhanced ability to fight conditional on a crisis. In Smith (1998b), a state’s decision to form an alliance conveys information to an aggressor about the disutility to fighting as well as serving as a commitment device for the state (see also Fearon 1997).

2.2 UN Authorizations and Coalitions of the Willing

The framework for UN authorizations of military force is contained Chapter VII of the UN Charter. Articles 39-42 of the Charter delegates to the UN Security Council the authority to determine whether a state’s action constitutes a threat to peace and, failing peaceful settlement of the crisis, to military action “to restore international peace and security.” The most prominent example is UN resolution 83, passed in June of 1950, which called on member states to provide military assistance to South Korea. (The Soviet Union, which could have vetoed the resolution, was not present due to its boycott of the Security Council at the time.) Article 51 of the UN charter recognizes the “inherent right of individual or collective self-defense. Resolution 1373, which authorized US military operations in Afghanistan following the attacks of September 11, 2001, fit under this rubric – although these sometimes require a broad interpretation of what is meant by “self-defense.” Typically, the Security Council deadlocks on such rulings, leading to General Assembly resolutions that are usually ignored (Voeten 2005). Since the first Gulf War in 1991, the frequency with which member states have sought authorization in the UN has increased dramatically, although these often fail to avoid a veto from a permanent security council member.

The term “coalition of the willing” was first used by President Bill Clinton, but is most closely associated with George W. Bush, who used the term to refer to countries that supported the invasion of Iraq in 2003. Despite strong lobbying by the US culminating in Colin Powell’s famous appearance before the General Assembly in February 2003, the US was unable to generate sufficient support for a Security Council resolution issuing an ultimatum to Saddam Hussein, and withdrew the resolution. The coalition invasion that began in March of 2003 was subsequently described by Secretary-General Kofi Annan as illegal from the perspective of the UN charter.
3 A Model

3.1 Primitives and Solution Concept

The interaction on which we focus is between a “rogue state” \( R \) and \( n \) other countries, indexed by \( i = 1, \ldots, n \). The rogue state benefits from taking a provocative action \( a \in [0, A] \), but may also incur costs imposed by the international community as well. The nature of these costs is the subject of our current examination.

In particular, assume that each state \( i \) is characterized by a cost parameter \( t_i \), which scales the disutility to \( i \) of the rogue state’s action. We assume that the parameter \( t_i \) is drawn from a common-knowledge prior distribution with probability density function \( p(t) \) and cumulative distribution function \( P(t) \). For now, we place no restrictions on the distribution of cost parameter other than to assume that (a) it can take on both negative and positive values, and (b) \( P(0) < 0.5 \) – i.e., in expectation, more states are likely to suffer than benefit from the rogue state’s action.

We will consider three international environments associated with potential responses to the rogue state:

1. **Pacts (PA):** Non-rogue states can choose whether to join, at fixed cost, a coalition that will impose sanctions on the rogue state proportional to its action above and beyond costs imposed privately.

2. **Strong United Nations (UN):** Member states in the UN vote, via unanimity rule, on whether to authorize action against the rogue state. If the vote succeeds, states can choose to join a UN enforcing coalition that will impose sanctions on the rogue state as in the PA case. If the vote fails, a coalition does not form.

3. **Breakaway Coalitions of the Willing (CW):** Member states in the UN vote, via unanimity rule, on whether to authorize action against the rogue state. If the vote succeeds, states can choose to join a UN enforcing coalition that will impose sanctions on the rogue state as in the PA case. If the vote fails, states can choose whether to join a a breakaway coalition that will impose sanctions on the rogue state as in the PA case.

For notational convenience, we will refer to generic institution as \( I \in \{ PA, UN, CW \} \).
The game unfolds as follows: first, the international community coordinates on sanctions for the rogue state conditional on the institutional environment and the expected response of the rogue state to those sanctions. Second, the rogue state observes the choices made by the other states: in the Pact Model, the rogue state observes the number of participants in the coalition. In the strong UN model, the rogue state observes the outcome of the UN vote and the vote tally and, if authorization was granted, the number of participants in the enforcing coalition. In the breakaway coalition model, the rogue state observes the outcome of the UN vote and the vote tally, the number of members of the enforcing coalition if the vote was successful, and the number of members of the breakaway coalition if the vote was unsuccessful. Finally, the rogue state chooses a level of action $a$, and sanctions are realized, with the rogue state incurring the corresponding disutility.

Before proceeding, we make four comments. First, note that in the model, the rogue state moves second, and thus the nature of the “provocative” action is itself an equilibrium concept. As will become clear below, because the rogue state’s preferences are commonly known, the sequence of actions we consider may be thought of as a subgame to a fuller game in which the rogue state can choose whether to set these events into motion by taking an initial provocative action.

Second, our model of a “strong” United Nations is meant to capture an environment in which the United Nations is effective at discouraging any international action outside of the UN framework. Thus, the breakaway coalition model may be thought of, correspondingly, as an environment with a “weak” UN.

Third, to focus on the underlying informational mechanism, we abstract away from two features of the international environment that observers often highlight as relevant to international affairs. First, we model heterogeneity in the costs a rogue states’ actions impose on non-rogue states, and not in the size of the latter group. The implication of this assumption is that all non-rogue states are equally capable of imposing costs on the rogue state for its actions; whether they choose to or not depends on the relative disutility of the rogue state’s actions and strategic effects. Second, we abstract away from side payments across states seeking to build coalitions. In our model, coalitions of non-rogue states are driven directly by their expected effects on the rogue state and the direct cost of participation.

Fourth, in the model the UN votes by unanimity rule in both the UN and CW models. This is intended to capture, in reduced form, the voting rules in the UN Security Council: Article 24 of the
UN Charter gives the Security Council “primary responsibility for the maintenance of international peace and security;” and on the Council, any of the five permanent members (China, the US, the UK, Russia, and France) may veto a resolution. This, of course, blurs the distinction between the security council and the general assembly, the latter of which votes either by majority or qualified majority rule on resolutions that are non-binding toward member states.

We assume that the costs that the rogue state incurs are twofold: first, irrespective of international action, non-rogue states may impose private costs on the rogue state in proportion to the severity of the rogue state’s action. Second, members of a coalition (of the willing, UN enforcing, or breakaway) may impose additional costs on the rogue state. Let \( \alpha \in \mathbb{R}^+ \) be a parameter that scales the former, and \( \beta \in \mathbb{R}^+ \) be a parameter that scales the latter. Let \( w_i \) be a dummy variable equal to one if state \( i \) participates in the sanctioning regime, and 0 otherwise. The rogue state’s utility associated with action \( a \) is

\[
\begin{align*}
u_R(a; t, w) &= \gamma a - \left( \alpha \sum_{i=1}^{n} t_i + \beta \sum_{i=1}^{n} w_i \right) a^2
\end{align*}
\]  

where, \( \gamma \) is a (commonly-known) parameter scaling the benefit associated with the provocative action, \( t \) is the vector of other states’ types, and \( w \) is the vector of other states’ actions. That the private costs imposed on the rogue state by individual non-rogue states are proportional to the non-rogue states’ types (as described by the vector \( t \)) is meant to capture, in reduced form, the policy consequences brought about by the domestic political outcry that would accompany the rogue state’s actions.

State \( i \)’s utility is given by

\[
u_i(w_i, t_i, a) = -(t_i + \delta w_i)a,
\]

where \( \delta \) is the cost to the non-rogue state associated with joining a coalition. In words, a coalition member shoulders, in addition to the direct cost of the rogue state’s action, an additional cost associated with sanctioning the rogue state.

The solution concept we employ is symmetric perfect Bayesian equilibrium. This requires that (a) each player’s choices be sequentially rational given her beliefs at the time of choice and
other players’ strategies; (b) conditional on type, all non-rogue states play the same strategy; and (c) beliefs about other players’ types be consistent with prior beliefs, equilibrium strategies, and Bayes’ Rule on the path of play.

3.2 Equilibrium Properties

The rogue state’s decision problem is identical across institutional environments, with one exception: the information available to it changes. Let \( k = \sum_{i=1}^{n} w_i \) be the number of non-rogue states participating in the international action against the rogue state. The rogue state’s expected utility associated with action \( a \), given information \( I(I) \), is given by

\[
E[u_r(a; t, w)|k, I] = \gamma a - \left( \beta k + \alpha E \left[ \sum_{i=1}^{n} t_i | I(I) \right] \right) a^2. \tag{2}
\]

Differentiating with respect to \( a \) gives the rogue state’s first order condition:

\[
\gamma - 2 \left( \beta k + \alpha E \left[ \sum_{i=1}^{n} t_i | I(I) \right] \right) a = 0.
\]

Solving for \( a \) and noting that this critical point is a maximum if and only if

\[
S(k, I(I)) \equiv \beta k + \alpha E \left[ \sum_{i=1}^{n} t_i | I(I) \right] > 0 \tag{3}
\]

yields:

\[
a^*(S(k, I(I))) = \begin{cases} 
\gamma(2S(k, I(I)))^{-1} \text{ if } S(k, I(I)) > 0 \\
A \text{ otherwise.} 
\end{cases} \tag{4}
\]

The expected utilities to state \( i \) of not participating and participating in a coalition are, respectively,

\[
E[u_i(w_i = 0|, t_i, a)] = -t_i E[a|w_i = 0] \\
E[u_i(w_i = 1|, t_i, a)] = -(t_i + \delta) E[a|w_i = 1].
\]

Thus, the state participates in the coalition if and only if

\[
-(t_i + \delta) E[a|w_i = 1] \geq -t_i E[a|\omega_i = 0], \tag{5}
\]
or

\[ t_i(E[a|w_i = 0] - E[a|w_i = 1]) \geq \delta E[a|w_i = 1]. \] (6)

Inequality (6) shows that non-rogue state actions are monotonic in type: the left side of the inequality is increasing linearly with \( t_i \) while the right side is a constant; thus, (6) satisfied at equality uniquely identifies \( t_i = \hat{t} \), which we use to denote the type of the non-rogue state that is exactly indifferent between joining and not joining a coalition.

Note that if \( E[a|w_i = 0] - E[a|w_i = 1] > 0 \), then \( \hat{t} > 0 \) and for all \( i \) s.t. \( t_i > \hat{t} \), \( w_i^* = 1 \). However, if \( E[a|w_i = 0] - E[a|w_i = 1] < 0 \), then \( \hat{t} < 0 \) and for all \( i \) s.t. \( t_i < \hat{t} \), \( w_i^* = 1 \). In the former case, joining the coalition decreases \( a^*(\cdot) \), both through the direct sanction \( \beta \) and by signaling the presence of a high type and thereby increasing \( E[\sum_{i=1}^n t_i I(I)] \). In the latter case, the effect of the direct sanction is the same, but joining signals the presence of a low type, thereby decreasing \( E[\sum_{i=1}^n t_i I(I)] \). If the effect of the direct sanction dominates, then there cannot be an informative equilibrium in which \( w_i^* = 1 \) if \( t_i < \hat{t} < 0 \) – such a strategy profile would entail non-rogue states that strictly benefited from a higher rogue-state action \( a \) actively discouraging it. If the signaling effect dominates, then it is possible to have an informative equilibrium in which low types join the coalition in order to signal to the rogue state that they benefit from its actions and will be rewarding it \( (-\alpha t_i > 0) \) rather than sanctioning it \( (-\alpha t_i < 0) \). Given the substantive motivation of this paper, and the fact that it is perverse for states who wish to credibly communicate their support to the rogue state to do so by imposing a punishment on it \( (\beta > 0) \) that is both costly to themselves and contrary to their purpose (in that \( a^*(\cdot) \) is decreasing in term \( \beta k \)), for the remainder of the paper, we will focus solely on the equilibrium in which \( w_i^* = 1 \) if and only if \( t_i \geq \hat{t} > 0 \).

The preceding discussion establishes the following result:

**Lemma 1** In equilibrium, states’ decisions to join the pact or the enforcing coalition under (strong) UN are weakly increasing in type and condition (6) holding at equality identifies the lowest type that joins.

The following lemma describes the relationship between the actions of the international community and the rogue state:

**Lemma 2** In any informative equilibrium, the rogue state’s action, characterized by (8), is
1. (weakly) decreasing in $S(k, I(\mathcal{I}))$ and thus in $k$; and

2. (weakly) convex in $S(k, I(\mathcal{I}))$.

That the rogue state’s action is decreasing in $k$ is intuitive: a higher $k$ indicates that a greater proportion of non-rogue states come from the high-end of the cost distribution, and that a high $a$ is likely to involve a substantial downside. If, however, the rogue state’s optimal action is the maximal one, $A$, then an increase in $k$ may have no effect – thus the relationship between $k$ and $a$ is weak. The marginal effect of a change in $k$ is itself decreasing because of the convexity of the rogue state’s costs: an increase in $k$ has a larger effect when $k$ is low than when it is high.

3.3 Comparing Institutions

Complete equilibrium characterizations of PA and UN, including equilibrium beliefs, are provided in the Appendix. before proceeding with the analysis of some of the equilibrium properties of those institutions, note that $S^{PA}$ is a specification of equation (3) for the pact environment; $S^{UN}$, which is the corresponding quantity for the UN environment, however, depends on an additional factor: the voting profile of the UN member-states. Let $v_i \in \{0, 1\}$ be state $i$’s vote against (0) or for (1) authorizing an enforcing coalition, where authorization requires a unanimous vote in favor. Define $v = \sum_{i=1}^{n} v_i$; then:

$$S^{UN}(v, k) = \begin{cases} \beta k + \alpha (kE[t_i|v_i = 1] + (n - k)E[t_i|v_i = 1, w_i = 0]) & \text{if } v = n \\ \alpha (vE[t_i|v_i = 1] + (n - v)E[t_i|v_i = 0]) & \text{if } v < n \end{cases}$$

In the remainder of this section, we explore several key features of the PA and UN equilibria. First, we compare rogue and non-rogue state behavior in each institution with behavior in an “anarchic baseline” (AB). Next, we compare behavior in the two institutional environments.

In the anarchic baseline, individual states may impose private ($\alpha$) costs to sanction a rogue state, but not the costs associated with a coalition. Note that this is equivalent to constraining the PA model to $\beta = 0$ and $k = 0$. The anarchic baseline should not be taken as a realistic description of anarchy. Rather, it is the normative benchmark against which, from an ex ante perspective, to which we would want to compare any alternative institutional arrangements.
First, we compare the ex ante expected action of the rogue state in the PA and UN equilibria with the expected action under AB. The following lemma characterizes the rogue state’s behavior in the anarchic baseline:

**Lemma 3** In the anarchic baseline, the rogue state chooses $a^*(S_{AB}) = \gamma \frac{\alpha n E[t_i]}{25}$, where $S_{AB} = \alpha n E[t_i]$.

Intuitively, in the absence of international action against the rogue state, the rogue state can learn nothing about the underlying type distribution of the non-rogue states, and, additionally, is not deterred by the direct consequences of coalition-driven sanctions. Consequently, it chooses a policy commensurate with its prior beliefs about the other states’ types.

Lemma 3 permits us to formulate the following result:

**Proposition 1** For sufficiently low values of $\beta$, the equilibrium expected rogue state action is higher in the Pact and in the Strong United Nations equilibria than in the anarchic baseline.

Informally, the rogue state’s action is a decreasing, convex function of $S$, which is, in turn, a function of the number of coalition members $k$. When $\beta$ is small, the ex ante expected marginal cost to the rogue state imposed by other states in the pact environment approaches the expected marginal cost in the anarchic baseline. By the convexity of the rogue state’s action, its expected action (the expectation taken over all possible draws of $k$) is higher than its action evaluated at the expected number of draws, which by the above argument approaches the action taken in the anarchic environment.

The above argument holds for the expected action taken by the rogue state ex ante – that is, before the other states know their own types. A useful question concerns the expected rogue state action from the perspective of a non-rogue state that knows its own type, and thus, by extension, its own expected welfare.

**Proposition 2** If $a^*(S_{AB}) > E[a^*(S_{PA}(k, \hat{t}_{PA}))|w_i = 0]$, then all $t_i < 0$ prefer AB to PA and all $t_i \geq 0$ prefer PA to AB. If $a^*(S_{AB}) < E[a^*(S_{PA}(k, \hat{t}_{PA}))|w_i = 0]$, then:

(a) if $(a^*(S_{AB}) - E[a^*(S_{PA}(k, \hat{t}_{PA}))|w_i = 1]) > 0$, then there is a type $\tilde{t} > \hat{t}_{PA}$ such that all $t_i \geq \tilde{t}$ and all $t_i < 0$ prefer PA to AB, and all other types $t_i \in [0, \tilde{t})$ prefer AB to PA; and
(b) if $(a^*(S^{AB}) - E[a^*(S^{PA}(k, i^{PA}))|w_i = 1]) < 0$, then $t_i < 0$ prefer PA to AB and $t_i > 0$ prefer AB to PA.

When the equilibrium rogue state action in AB is greater than its expected action in PA given at least one state is certain to join the pact, but less than its expected action given at least one state is certain not to join, the state’s preferences over institutions are not monotonic. In particular, some of the types that join the pact in equilibrium actually prefer AB. In this sense, thinking of them as “willing” participants is somewhat misleading, since they prefer an institutional framework in which coalitions are not possible, even though once a coalition is possible, they are better off joining it than not. They prefer AB even though they anticipate that the rogue state’s action will be more aggressive in AB because the cost they will bear from membership in the coalition is greater than the benefits they receive from the difference in the rogue state’s actions across institutions.

States that do not join the pact in equilibrium anticipate that the rogue state will be more aggressive in expectation and conditional on their own non-participation, in PA and in AB. Accordingly, such states prefer PA if they benefit from the rogue state’s action and AB otherwise. Thus, the lowest and the highest types prefer PA, which those in the middle prefer AB.

In the anarchic baseline, a decision by a non-rogue state not to participate conveys no information to the rogue state regarding the type of the former. By contrast, under PA or UN, the same act of non-participation is informative, and would make the non-rogue state strictly worse off in expectation than under the anarchic baseline.

The following proposition compares rogue and non-rogue state behaviors in the PA and UN environments:

**Proposition 3** (1) Holding fixed the type-threshold for coalition membership between the pact and the UN’s enforcing coalition, the UN’s enforcing coalition is more effective in suppressing the rogue state’s action in expectation. (2) In equilibrium, the expected size of the enforcing coalition that forms as a consequence of UN authorization is smaller than the expected size of a pact.

Note that the comparison of the PA and the UN suggests a fundamental tradeoff. The benefit of the UN system is that, conditional on receiving authorization, the expected benefit of international organization to non-rogue states is higher in expectation. However, employing such
a system runs the risk of authorization failures, which do not come into play under pacts. Holding fixed the informational effect, the downside risk is highest for the highest types. By contrast, the coalition of the willing has a weaker expected effect than that of a UN enforcing coalition conditional on authorization, but strongly affected states always have a recourse to them. If one considers the United States to be such a strongly affected state, then this intuition accords with former Ambassador Bolton’s first argument as described in the Introduction. At the same time, pacts circumscribe the rogue state’s ability to learn and potentially exploit information regarding the type distribution of non-participants compared to the environment with a strong UN – consistent with Bolton’s second (informational) argument against the institution.

3.4 Weak UN with Breakaway Coalitions of the Willing

Suppose it is possible to form a break-away coalition after a failed attempt to obtain UN authorization.

Let $b_i \in \{0,1\}$ denote state $i$’s decision to participate ($b_i = 1$) or not to participate ($b_i = 0$) in a break-away coalition following a non-unanimous vote to obtain a UN authorization. Define

$$S_{CW}(v,k) = \begin{cases} 
S^{UN}(v,k) & \text{if } v = n \\
\beta k + \alpha(kE[t_i|b_i = 1] + (v-k)E[t_i|v_i = 1, b_i = 0] + (n-v)E[t_i|v_i = 0]) & \text{if } v < n 
\end{cases}$$

Note that if $v = n$, $k$ refers to the number of states that join the UN enforcement coalition, and if $v < n$, $k$ refers to the number of states that join CW.

**Proposition 4** There exists an informative equilibrium of the CW game in which

1. A non-rogue state $i$ votes in favor of authorizing the enforcing coalition if and only if $t_i \geq 0$.

2. A non-rogue state $i$ participates in the enforcing coalition, $w_i^* = 1$, if and only if

$$t_i \geq \hat{t}^{EC} = \frac{\delta E[a^*(SCW(n,k))|w_i = 1]}{E[a^*(SCW(n,k))|w_i = 0] - E[a^*(SCW(n,k))|w_i = 1]}$$

if $v = n$ and

$$t_i \geq \hat{t}^{CW} = \frac{\delta E[a^*(SCW(v,k))|b_i = 1]}{E[a^*(SCW(v,k))|b_i = 0] - E[a^*(SCW(v,k))|b_i = 1]}$$

if $v < n$. 

13
(3) The rogue state chooses

\[ a^*(S_{CW}(v,k)) = \begin{cases} 
\gamma (2S_{CW}(v,k))^{-1} & \text{if } S_{CW}(v,k) > 0 \\
A & \text{otherwise}
\end{cases} \]

(4) The rogue state’s beliefs are

\[
E[t_i|v_i = w_i = 1] = E[t_i|t_i \geq t^{UN}]
\]
\[
E[t_i|v_i = 1, w_i = 0] = E[t_i|0 \leq t_i < t^{UN}]
\]
\[
E[t_i|v_i = 1, b_i = 1] = E[t_i|t_i \geq t^{CW}]
\]
\[
E[t_i|v_i = 1, b_i = 0] = E[t_i|0 \leq t_i < t^{CW}]
\]
\[
E[t_i|v_i = 0] = E[t_i|t_i < 0].
\]

(5) And the equilibrium non-rogue states’ expectation of the rogue state action given \( w_i \) are

\[
E[a^*(S_{CW}(n,k))|w_i = 1] = \sum_{k_{-i}=0}^{n-1} \binom{n-1}{k_{-i}} \frac{\gamma(1 - P(t^{UN}))^{k_{-i}} P(t^{UN})^{n-1-k_{-i}}}{2S_{CW}(v=1,k_{-i}+1)}
\]
\[
E[a^*(S_{CW}(n,k))|w_i = 0] = \sum_{k_{-i}=0}^{n-1} \binom{n-1}{k_{-i}} \frac{\gamma(1 - P(t^{UN}))^{k_{-i}} P(t^{UN})^{n-1-k_{-i}}}{2S_{CW}(v=1,k_{-i})}
\]
\[
E[a^*(S_{CW}(v,k))|b_i = 1] = \sum_{k_{-i}=0}^{v-1} \binom{v-1}{k_{-i}} \frac{\gamma(1 - P(t^{CW}))^{k_{-i}} P(t^{CW})^{v-1-k_{-i}}}{2S_{CW}(v,k_{-i}+1)(1-P(0))^{v-1}}
\]
\[
E[a^*(S_{CW}(v,k))|b_i = 0] = \sum_{k_{-i}=0}^{v-1} \binom{v-1}{k_{-i}} \frac{\gamma(1 - P(t^{CW}))^{k_{-i}} P(t^{CW})^{v-1-k_{-i}}}{2S_{CW}(v,k_{-i})(1-P(0))^{v-1}}
\]

Our final set of results considers some of the properties of the (weak) UN.

**Proposition 5** If a sufficiently large number of states vote against the authorization, then the minimum type that is willing to join the break-away coalition is lower than the minimum type that would join PA.

**Corollary 1** For sufficiently “bad” defeats of authorization requests, there are types of states that would join a break-away coalition that would have joined neither the enforcing coalition (under the strong UN) had authorization succeeded, nor PA.
Proposition 6  

1. The marginal value of joining a break-away coalition is increasing in the 
number of votes against authorization.

2. For any non-unanimous UN vote and given a sufficiently low technological advantage of joint 
action $\beta$, there exist types that have a higher expected welfare sticking with a strong UN than 
with the possibility of a break-away coalition (i.e., prefer that possibility of CW not exist) that 
will, nonetheless, join the break-away coalition, given it exists.

That a non-rogue state might, ex ante, prefer a strong UN, but would nonetheless join a 
breakaway coalition in the event of a UN authorization failure is consistent with the position of 
British Prime Minister Tony Blair as a reluctant coalition member in the 2003 Iraq war. At the 
time, Blair expressed a preference for going through the UN to disarm Iraq, but nonetheless felt 
compelled to join Bush’s coalition when a new resolution more severe than the one in place at the 
time (1441) was a non-starter in the Security Council.

References


Disputes.” American Political Science Review 88: 577-592.

Fearon, James D. 1997. “Signaling Foreign Policy Interests: Tying Hands versus Sinking 


Resolution 38: 270-297.


flict Resolution 48: 459-486.

Appendix

Proof of Lemma 2

Suppose \( a^*(S(k, I)) < A \). Then \( S(k, I) > 0, \frac{\partial a^*(\cdot)}{\partial S} = -\gamma (S(k, I))^{-2} < 0, \) and \( \frac{\partial^2 a^*(\cdot)}{\partial S^2} = \gamma (S(k, I))^{-3} > 0 \). From the monotonicity of \( w^*_i \) in \( t_i \), \( S(k, I) \) must be a positive linear function of \( k \):

\[
\beta k + \alpha k E[t_i|w^*_i = 1] + \alpha (n-k) E[t_i|w^*_i = 0],
\]

where \( \beta > 0, \alpha > 0, \) and \( E[t_i|w^*_i = 1] > E[t_i|w^*_i = 0] \) given \( w^*_i = 1 \) if and only if \( t_i \geq \hat{t} \). Thus, \( \frac{\partial a^*(\cdot)}{\partial S} \) is decreasing in \( k \) and \( \frac{\partial^2 a^*(\cdot)}{\partial S^2} \) is decreasing in \( k \) but always remains positive.

Pacts: Equilibrium Characterization

The following lemma characterizes equilibrium behavior under PA:

Lemma 4 The following profile of strategies and beliefs is an equilibrium to the Pact game:

1. A non-rogue state \( i \) participates in a coalition, \( w^*_i = 1 \), if and only if

\[
t_i \geq \hat{t}^{PA} = \frac{\delta E[a^*(\cdot)|w_i = 1]}{E[a^*(\cdot)|w_i = 0] - E[a^*(\cdot)|w_i = 1]},
\]
2. The rogue state chooses

\[
a^*(S^{PA}(k)) = \begin{cases} 
\gamma(2S^{PA}(k))^{-1} & \text{if } S^{PA}(k) > 0 \\
A & \text{otherwise}
\end{cases}
\]  

(8)

where

\[
S^{PA}(k) = \beta k + \alpha k E[t_i|w_i^* = 1] + \alpha(n-k)E[t_i|w_i^* = 0].
\]

3. The rogue state’s beliefs are given by

\[
E[t_i|w_i^* = 1] = E[t_i|t_i \geq \hat{t}^{PA}]
\]

\[
E[t_i|w_i^* = 0] = E[t_i|t_i < \hat{t}^{PA}].
\]

4. And the equilibrium non-rogue states’ expectation of the rogue state action is

\[
E[a|w_i = 1] = \sum_{k-i=0}^{n-1} \binom{n-1}{k-i} \frac{\gamma(1-P(\hat{t}))^{k-i} P(\hat{t})^{n-1-k-i}}{2S^{PA}(k-i+1)}
\]

\[
E[a|w_i = 0] = \sum_{k-i=0}^{n-1} \binom{n-1}{k-i} \frac{\gamma(1-P(\hat{t}))^{k-i} P(\hat{t})^{n-1-k-i}}{2S^{PA}(k-i)}.
\]

Proof: Part 1 follows from (6); part 2 from (8) and the specification of \(S(k,I)\) introduced in (3); part 3 follows from part 1, and part 4 from (8), monotonicity of \(w_i\) in \(t_i\), and basic combinatorics.

Equilibrium Characterization: Strong United Nations

Let \(v_i, v, \) and \(S^{UN}(v,k)\) be as defined in the main body of the paper. Then:

**Lemma 5** There exists an informative equilibrium of the Strong UN game in which:

(1) A non-rogue state \(i\) votes in favor of authorizing the enforcing coalition if and only if \(t_i \geq 0\).

(2) Given \(v = n\), a non-rogue state \(i\) participates in the enforcing coalition, \(w_i^* = 1\), if and only if

\[
t_i \geq \hat{t}^{UN} = \frac{\delta E[a^*(S^{UN}(n,k))|w_i = 1]}{E[a^*(S^{UN}(n,k))|w_i = 0] - E[a^*(S^{UN}(n,k))|w_i = 1]}.
\]
(3) The rogue state chooses

\[ a^*(S^{UN}(v,k)) = \begin{cases} \gamma (2S^{UN}(v,k))^{-1} & \text{if } S^{UN}(v,k) > 0 \\ A & \text{otherwise} \end{cases} \]

(4) The rogue state’s beliefs are

\[ E[t_i|v_i = w_i = 1] = E[t_i|t_i \geq \hat{t}^{UN}] \]
\[ E[t_i|v_i = 1, w_i = 0] = E[t_i|0 \leq t_i < \hat{t}^{UN}] \]
\[ E[t_i|v_i = 1] = E[t_i|t_i \geq 0] \]
\[ E[t_i|v_i = 0] = E[t_i|t_i < 0]. \]

(5) And the equilibrium non-rogue states’ expectation of the rogue state action given \( w_i \) are

\[ E[a^*(S^{UN}(n,k))|w_i = 1] = \sum_{k_{-i}=0}^{n-1} \binom{n-1}{k_{-i}} \frac{\gamma(1 - P(\hat{t}^{UN}))^{k_{-i}} P(\hat{t}^{UN})^{n-1-k_{-i}}}{2S^{UN}(v = 1, k_{-i} + 1)} \]
\[ E[a^*(S^{UN}(n,k))|w_i = 0] = \sum_{k_{-i}=0}^{n-1} \binom{n-1}{k_{-i}} \frac{\gamma(1 - P(\hat{t}^{UN}))^{k_{-i}} P(\hat{t}^{UN})^{n-1-k_{-i}}}{2S^{UN}(v = 1, k_{-i})}. \]

Proof:

Suppose \( v = n \), then the subsequent subgame is analogous to the PA game, substituting \( S^{UN}(n,k) \) for \( S^{PA}(k) \) and censoring the distribution of \( t_i \) to reflect \( v_i = 1 \) \( \forall i \).

Suppose \( v < n \). Then the subsequent subgame is analogous to the AB, substituting \( S^{UN}(v,\cdot) \) for \( S^{AB}(\cdot) \) if \( S^{UN}(v,\cdot) > 0 \). If \( S^{UN}(v,\cdot) \leq 0 \), then \( a^*(S^{UN}(v,\cdot)) = A \). Thus, Proposition 4 and Lemma 3 establish parts 2, 3, 5, and part of 4.

We next construct the optimal decision rule for the states’ choices to vote in favor of the authorization \( (v_i = 1) \). There are three distinct intervals:
For $t_i \geq \hat{t}^{UN}$: state $i$ will vote in favor iff

$$\Pr(v_j = 1 \ \forall j \neq i) \left[ -(t_i + \delta)E[a^*(S^{UN}(n,k))|w_i = 1, v = n] + t_i E[a^*(S^{UN}(v,\cdot))|v = n - 1] \right] + \sum_{h=0}^{n-2} \Pr(\sum_{j \neq i} v_j = h) \left[ -t_i E[a^*(S^{UN}(v,\cdot))|v = h + 1] + t_i E[a^*(S^{UN}(v,\cdot))|v = h] \right] \geq 0. \tag{9}$$

For $t_i \in (0, \hat{t}^{UN})$: state $i$ will vote in favor iff

$$\Pr(v_j = 1 \ \forall j \neq i) \left[ -(t_i + \delta)E[a^*(S^{UN}(n,k))|w_i = 0, v = n] + t_i E[a^*(S^{UN}(v,\cdot))|v = n - 1] \right] + \sum_{h=0}^{n-2} \Pr(\sum_{j \neq i} v_j = h) \left[ -t_i E[a^*(S^{UN}(v,\cdot))|v = h + 1] + t_i E[a^*(S^{UN}(v,\cdot))|v = h] \right] \geq 0. \tag{10}$$

For $t_i < 0$, the condition is the same as for $t_i \in (0, \hat{t}^{UN})$, but if it holds for $t_i > 0$, it must fail for $t_i < 0$ and vice versa.

Note that LHS of (10) is a multiplicative constant times $t_i$. Thus, if (10) holds for some $t_i > 0$, it holds for all $t_i \geq 0$, including $t_i = \hat{t}^{UN}$, and fails for all $t_i < 0$. Note also that if (10) is satisfied for $t_i = \hat{t}^{UN}$, then (9) must be satisfied for $t_i = \hat{t}^{UN}$ as well, and the LHS of (9) must be an increasing function of $t_i$. Thus (9) is satisfied for all $t_i \geq \hat{t}^{UN}$ if (10) is satisfied for some $t_i > 0$.

Finally, observe that $v_i = 1$ iff $t_i \geq 0$ implies part 4 of the proposition, which in turn determines $S^{UN}(v,k)$ and $a^*(S^{UN}(v,k))$, insuring that the LHS of (10) is a positive, increasing function of $t_i$. ■

**Proof of Lemma 3**

In the anarchic baseline, $\beta = 0$, and the costs imposed by the non-rogue states are purely a function of their types and not of any choices on their part. Thus, the rogue state learns nothing about their types, and $S^{AB} = \alpha n E[t_i]$. By assumption, $E[t_i] > 0$. Therefore, $S^{AB} > 0$, implying an interior choice of $a^*(S^{AB})$. Substituting into (8) gives the expression in the proposition. ■
Proof of Proposition 1

Let $\beta = 0$. First we show that a separating equilibrium exists under PA for $\beta = 0$, and then we show that the ex ante expected rogue state action is higher in that separating equilibrium than in the equilibrium of the anarchic baseline.

$$S_{\beta=0}^{PA}(k) = \alpha(kE[t_i|w_i = 1] + (n - k)E[t_i|w_i = 0]).$$

Suppose

$$E[a^*(S_{\beta=0}^{PA}(k))|w_i = 0] > E[a^*(S_{\beta=0}^{PA}(k))|w_i = 1].$$

Then from (7), there exists $i^{PA}_{\beta=0} > 0$ s.t. $w_i = 1$ for all $t_i \geq i^{PA}_{\beta=0}$, and from Lemma 2 and Proposition 2, there is a separating equilibrium consistent with the supposition.

By the law of iterated expectations, $E[S_{\beta=0}^{PA}(k)] = \alpha nE[t_i] = S^{AB}$. Let $k$ be the smallest integer such that $S_{\beta=0}^{PA}(k) > 0$. From Lemma 1, $a^*(S_{\beta=0}^{PA}(k))$ is strictly decreasing and convex for all $S_{\beta=0}^{PA}(k) > 0$, and for $S_{\beta=0}^{PA}(k)$, $a^*(S_{\beta=0}^{PA}(k)) = A$. Pr($a^*(S_{\beta=0}^{PA}(k)) = A$) = Pr($k < k$). By Jensen’s inequality,

$$E[a^*(S_{\beta=0}^{PA}(k))] > a^*E[(S_{\beta=0}^{PA}(k))] = a^*(S^{AB}).$$

By continuity, the inequality holds for $\beta = \epsilon$ for some $\epsilon > 0$.

An equivalent proof establishes the result comparing UN with AB. ■

Proof of Proposition 2

Consider $t_i \geq i^{PA}$. From Proposition 4, $w_i = 1$ for $t_i$ in PA. Thus, $t_i$ prefers PA to AB iff

$$E[u_i(w_i = 1, t_i, a^*(S^{PA}(k,i^{PA})))|w_i = 1] \geq E[u_i(w_i = 0, t_i, a^*(S^{AB})))].$$

This is equivalent to

$$t_i(a^*(S^{AB}) - E[a^*(S^{PA}(k,i^{PA}))|w_i = 1]) \geq \delta E[a^*(S^{PA}(k,i^{PA}))|w_i = 1].$$

(11)
Observe that there exists \( t_i \) s.t. this inequality holds iff

\[
(a^*(S^{AB}) - E[a^*(S^{PA}(k,i^{PA}))|w_i = 1]) > 0. \tag{12}
\]

Consider \( t_i < \hat{t}^{PA} \). From Proposition 4, \( w_i = 0 \) for \( t_i \) in PA. Thus, \( t_i \) prefers PA to AB iff

\[
E[u_i(w_i = 0, t_i, a^*(S^{PA}(k,i^{PA})))|w_i = 0] \geq E[u_i(w_i = 0, t_i, a^*(S^{AB}))].
\]

This is equivalent to

\[
t_i(a^*(S^{AB}) - E[a^*(S^{PA}(k,i^{PA}))|w_i = 0]) \geq 0.
\]

Therefore, if

\[
a^*(S^{AB}) - E[a^*(S^{PA}(k,i^{PA}))|w_i = 0]) > 0, \tag{13}
\]

then all \( t_i \in [0, \hat{t}^{PA}) \) prefer PA and all \( t_i < 0 \) prefer AB; if the sign of the inequality (13 is reversed, but inequality is strict, then all \( t_i \leq 0 \) prefer PA and all \( t_i \in (0, \hat{t}^{PA}) \) prefer AB.

Because

\[
E[a^*(S^{PA}(k,i^{PA}))|w_i = 1] < E[a^*(S^{PA}(k,i^{PA}))|w_i = 0],
\]

there are three possible cases to consider: (1) both conditions (12) and (13) hold; (2) condition (12) holds, but condition (13) fails; and (3) both conditions (12) and (13) fail.

Case (1). Let \( \tilde{t} \) be \( t_i \) that satisfies (11) at equality. Then, from (13), (6), and (11), \( \tilde{t} < \hat{t}^{PA} \), and thus \( t_i < 0 \) prefers AB and \( t_i \geq 0 \) prefers PA.

Case (2). If the case conditions hold, then from (13) and (6), \( \tilde{t} > \hat{t}^{PA} \) and thus, \( t_i < 0 \) and \( t_i \geq \tilde{t} \) prefer PA and \( t_i \in (0, \tilde{t}) \) prefer AB.

Case (3). If the case conditions hold, then \( t_i < 0 \) prefer PA and \( t_i > 0 \) prefer AB. \( \blacksquare \)

**Proof of Proposition 3**

Equation (8) determines \( a \). \( S^{PA}(k) \) and \( S^{UN}(n,k) \) have the same form, but \( E[t_i|w_i^* = 0], E[t_i|w_i^* = 1] \) in PA and \( E[t_i|v_i^* = 1, w_i^* = 0], E[t_i|v_i^* = w_i^* = 1] \) in UN are different. Because \( v_i^* = 1 \) implies \( t_i \geq 0, E[t_i|t_i < \tilde{t}] < E[t_i|0 \leq t_i < \tilde{t}] \), holding constant \( \hat{t} \).
Imposing $i^{PA} = i^{UN} = \hat{i}$ (which is not equilibrium) has the (direct) effect of making $S^{UN}(n, k) > S^{PA}(k)$ for $k < n$ (and $S^{UN}(n, n) = S^{PA}(n)$). Thus, $a^*(S^{UN}(n, k)) < a^*(S^{PA}(k))$, establishing part 1.

The indirect effects are two-fold. First, it raises $E[u_i(w_i = 0, t_i, a^*(S^{UN}(n, k)))|w_i = 0]$, which, from (7), increases $\hat{i}^{UN}$ relative to $\hat{i}^{PA}$. Second, because $a^*(S)$ is decreasing and convex (from lemma 2), $S^{UN}(n, k) > S^{PA}(k)$ implies the marginal benefit to $i$ of $w_i = 1$ is lower under UN than under PA, which also increases $\hat{i}^{UN}$ relative to $\hat{i}^{PA}$. Thus, in equilibrium $\hat{i}^{UN} > \hat{i}^{PA}$. ■

Proof of Proposition 7

The subgame following $v = n$ is the same as in the UN game. The subgames following each $v < n$ are analogous to the PA game, substituting $S^{CW}(v, k)$ for $S^{PA}(k)$, censoring the distribution of $t_i$ to reflect $v_i = 1$ for $v$ of the players and to reflect $v_i = 0$ for $n - v$ of the players. Thus, Propositions 1 and 2 establish parts 1-3 and part of part 4; given $v_i = 1$ iff $t_i \geq 0$, they also establish the remainder of part 4 and part 5. To see that in equilibrium $v_i = 1$ iff $t_i \geq 0$, observe that $t_i \geq 0$ prefers to vote for authorization iff

$$\Pr(v_j = 1 \forall j \neq i)[-(t_i + \delta)E[a^*(S^{UN}(n, k))|w_i = 1, v = n] + t_i E[a^*(S^{CW}(n - 1, k)|v = n - 1, w_i)]$$
$$+ \sum_{h=0}^{n-2} \Pr(\sum_{j \neq i} v_j = h)[-t_i E[a^*(S^{CW}(h + 1, k))|v = h + 1, w_i] + t_i E[a^*(S^{CW}(h, k))|v = h, w_i]] \geq 0.$$ 

Note that this condition is analogous to (9). Similarly, $t \in (0, \hat{i}^{UN})$ prefers to vote for authorization iff

$$\Pr(v_j = 1 \forall j \neq i)[-t_i E[a^*(S^{UN}(n, k))|w_i = 0, v = n] + t_i E[a^*(S^{CW}(n - 1, k)|v = n - 1, w_i)]$$
$$+ \sum_{h=0}^{n-2} \Pr(\sum_{j \neq i} v_j = h)[-t_i E[a^*(S^{CW}(h + 1, k))|v = h + 1, w_i] + t_i E[a^*(S^{CW}(h, k))|v = h, w_i]] \geq 0,$$

which is an analogue of condition (10).

Note that the last condition can still be factored by $t_i$, and all the logic of the proof of Proposition applies. ■
Proof of Proposition 8

To be added...

Proof of Proposition 9

(1) For \( v < n \),

\[ SCW(v, k + 1) - SCW(v, k) = \beta + \alpha(E[t_i|t_i \geq \hat{t}^{CW}] - E[t_i|0 \leq t_i < \hat{t}^{CW}]) \]

Note that this difference is constant regardless of \( v \) and \( k \). From the definition of \( SCW(v, k) \) and \( E[t_i|0 \leq t_i < \hat{t}^{CW}] > E[t_i|t_i \geq \hat{t}^{CW}] \), \( SCW(v, k) \) is increasing in \( v \). From Lemma 2, \( a^* \) is decreasing and convex in \( SCW(v, k) \). Thus at higher values of \( v \), the effect of a \( \beta + \alpha(E[t_i|t_i \geq \hat{t}^{CW}] - E[t_i|0 \leq t_i < \hat{t}^{CW}) \) increase in \( SCW \) on \( E[a^*(\cdot)] \) is smaller.

(2) The second part of the proposition states that for any \( v < n \), for sufficiently small \( \beta \), \( \exists t_i \geq \hat{t}^{CW}(v) \) s.t.

\[-(t_i + \delta)E[a^*(SCW(v, k))|b_i = 1] < -t_i a^*(S^{UN}(v, \cdot)).\]

From the informativeness of state \( i \)'s decision not to join a breakaway coalition,

\[ E[a^*(SCW(v, k))|b_i = 0] > E[a^*(SCW(v, k))]. \]

From Lemma 2 and Jensen's Inequality, \( E[a^*(SCW(v, k))] > a^*(E[SCW(v, k)]) \). Let \( \beta = 0 \). Then \( a^*(E[SCW(v, k)]) = a^*(S^{UN}(v, \cdot)) \). Thus, for \( \beta = 0 \),

\[-t_i E[a^*(SCW(v, k))|b_i = 0] < -t_i a^*(S^{UN}(v, \cdot)).\]

Let \( t_i = \hat{t}^{CW} \). Then

\[-(\hat{t}^{CW} + \delta)E[a^*(SCW(v, k))|b_i = 1] = -\hat{t}^{CW} E[a^*(SCW(v, k))|b_i = 0],\]

and for \( \beta = 0 \),

\[-(\hat{t}^{CW} + \delta)E[a^*(SCW(v, k))|b_i = 1] < -\hat{t}^{CW} a^*(S^{UN}(v, \cdot)).\]
By continuity, this inequality holds for $\beta = \varepsilon$, for $\varepsilon$ sufficiently small.