Credibility in Compliance and Punishment: Leader Specific Punishments and Credibility.*

Fiona McGillivray
Department of Politics
New York University
726 Broadway, NY, NY 10003
Fiona.McGillivray@nyu.edu

Alastair Smith
Department of Politics
New York University
726 Broadway, NY, NY 10003
Alastair.Smith@nyu.edu

October, 2004

ABSTRACT

The ability of nation A to compel nation B to grant it concessions depends upon the credibility with which A can commit to punish B for non-compliance. Discarding traditional unitary actor approaches, we assume policy in each nation is set by mortal political leaders and model the compliance/punishment relation between A and B in a noisy infinitely repeated setting. If nations utilize leader specific punishment, that is target policies against leaders rather than the nations they represent, then leaders improve the credibly of their threats to punish non-compliance since citizens have incentives to depose leader to restore national integrity.

*An earlier version of this paper was presented at the Center for International Studies Conference on Compliance and International Law, University of Southern California, April 30 – May 1, 2004. We thank the participants at that conference for their comments. We also thanks George Downs, Bruce Bueno de Mesquita and several anonymous referees for useful comments.

The limits of international cooperation and the limits to which nation A can compel another nation, B, to shifts its policies to A's liking depend upon the extent to which nation A can credibly commit to punish nation B for non-compliance. We combine recent theoretical insights of Leader Specific Punishments (McGillivray and Smith 2000) and design of international organizations (Rosendorff 2004; Rosendorff and Milner 2001) to examine how much cooperation and compliance can be enforced in a noisy environment of imperfect information. We show that when domestic political institutions are such that the domestic replacement of leaders is easy, as it is in democratic nations, punishment strategies targeted against individual leaders increase the credibility of punishment and compliance and hence increase the range of conditions under which A can credibly compel nation B.

Traditional liberal approaches to international cooperation emphasis the role of reciprocal punishments in generating a commitment to an agreement (Axelrod and Keohane 1986; Keohane 1984). Although these approaches often attempt to integrate domestic politics, most often in the form of shaping national preferences, the liberal paradigm typically analyses interstate relations from the perspective of unitary actor nations. Here we shift the unit of analysis from the nation state to individual leaders. Building on McGillivray and Smith's (2000) concept of Leader Specific Punishments (LSP), we assess the impact of having sanctions and punishments targeted against individual leaders rather than the nations they represent. We derive how the limits on compliance and cooperation depend upon domestic political institutions and show how the pattern of interactions between states depend upon domestic political institutions and leader turnover. Leader specific punishments enable higher levels of cooperation and compliance to be credibly implemented.

McGillivray and Smith (2000, 2004a,b) introduced the concept of leader specific punishment (LSP) in the context of the infinitely repeated prisoners' dilemma game to show that punishments targeted against individual leaders rather than nations per se improve the range of conditions under which cooperation is possible provided domestic political institutions allow for the relatively easy replacement of political leaders. The basis idea behind LSP is that nation A only punishes nation B for cheating or renegaging an agreement for as long as the leader in nation B who implemented the cheating remains in power. In this context, if β is the leader of nation B then nation B's reputation or integrity really belongs to leader β. Similarly, nation A's reputation is associated with its leader α. Since under LSP, A holds a grudge against β rather than B, the replacement of leader β provides the opportunity to normalize relations between states.

Recent accounts of US relations with Yugoslavia and Iraq illustrate the leader specific nature of US security and sanctions policy. US policy explicitly targeted sanctions against the leadership of Slobodan Milosevic of Yugoslavia and Saddam Hussein of Iraq. The removal of these leaders ended economic sanctions and reinvigorated relations. McGillivray and Stam (2004) show a systematic relationship between leader turnover and termination of economic sanctions. Since leader turnover restores cooperative relations and ends punishment, LSP strategies provide the citizens of a nation with an incentive to depose any leader who incurs the ire of foreign states. By removing such a leader the citizens can end the punishment phase and normalize relations. This provides citizens with a motive to replace leaders following the breakdown of cooperation. Since political leaders want to retain office, leaders who are easily replaced avoid actions that endanger international cooperation. Through this mechanism the level of cooperation obtainable between nations with easily replaced leaders is much higher than the level of cooperation possible between unitary actor states.[1]

While insightful, the McGillivray and Smith (2000) model considers only the simple

[1] In the context of cooperation between ethnic groups, Fearon and Laitin (1996) show that intragroup punishment can enhance the prospects for intergroup cooperation.

case of the prisoners' dilemma with its binary choice to cooperate or defect in a world of perfect information. Here we generalize the study of LSP by considering an infinitely repeated compliance/punishment game of policy disagreements between nations A and B in a noisy environment with mortal (i.e. not infinitely lived) leaders. The basic interaction between nations A and B is as follows. At time t, β, the leader of nation B, picks a policy x_{Bt} between 0 and 1. We assume that nation B (or at least the political elite of that state) want to set the policy as high as possible. In contrast, nation A wants policy set as low as possible. Given policy choice x_{Bt} nation B receives a payoff of x_{Bt}, while A's payoff is $(1 - x_{Bt})$. We examine the extent to which nation A can compel nation B to maintain policy below some specified threshold, h, through the threat of sanctioning non-compliance.

The nature of the policy disagreement in the compliance/punishment game can be thought of in two ways. First, the game speaks directly to attempts by nation A to compel nation B to alter its policies to A's wishes. Examples of such interactions are common place. For instance the US pressures nations to improve their human rights records, assist in the war on terror, help stem the flow of illegal drugs and uphold international copyright laws. In these examples nation A requests that nation B sets policy away from what B would ideally choose, $x_{Bt} = 1$, by threatening sanctions as punishment if B's policy is observed above the threshold h. A second alternative interpretation of the model is that it represents half of a cooperative arrangement. Trade agreements can be thought of in this context. The threshold h represents the limit of the tariffs B can impose in A's exports. The other (un-modeled) half of the agreement specifies A's concessions to B. Although the results of the model have implications for cooperative agreements, for ease of language we focus on the first compliance/punishment case.

The world is a noisy place. As such it is not always possible to directly observe nation B's policy. For example, nation A can not perfectly monitor nation B's human rights record. When nation B decides how many people to torture, it can not be certain how many of these instances will be reported in nation A. While nation B's policy is x_{Bt}, nation A observes $y_{Bt} = x_{Bt} + \varepsilon_{Bt}$, where ε_{Bt} is a stochastic error with iid distribution $F(\varepsilon_{Bt})$. For convenience we assume $F(\varepsilon_{it})$ is continuous, concave and twice differentiable with full support. We restrict attention to the exponential distribution. See Downs and Rocke (1995), Green and Porter (1984) and Porter (1983) for related models.

The probability, p, that nation B is found in breach of the 'rules' defined by the threshold h increases as the policy choice x_{Bt} increases. Specifically, $p = \Pr(y_{Bt} > h) = \Pr(x_{Bt} + \varepsilon_{Bt} > h) = 1 - F(h - x_{Bt})$. The higher nation B sets its policy choice, the greater the rewards from the policy but the more likely it is to be caught breaking the rules. This is one of the fundamental trade-offs nations face when setting policy (Downs and Rocke 1995).

When leaders set policy one of their essential consideration is the consequence of being 'caught'. If there is no effective punishment for being caught then nations pay little heed to the rules. In contrast, if detected violations result in harsh punishments then leaders are more likely to set policy low so as to reduce the risk of being caught. We assume nation A can apply sanctions to nation B. These sanctions impose a cost c_B on nation B and cost nation A c_A to send. The key insight of leader specific punishments is that in political systems where leader replacement is relatively easy citizens will depose leaders who in some sense do not 'play by the rules.' We say a leader has integrity but loses it if she does not play by the rules. In the current context, α maintains her integrity by sanctioning β when ever appropriate. Leader β loses her integrity if her policy is observed to cross the threshold, $y_{Bt} > h$. Once in breach of such a threshold, β's integrity depends upon her willingness to apologize by making additional policy concessions. The willingness of the citizens to depose their leader to restore national integrity magnifies the effective size of the sanction that a leader faces by endangering a leader's hold on office. Domestic institutions and leader turnover shape the compliance/punishment relationship.

In international relations most formal models consider repeated relations between nations or leaders as interactions between infinitely lived agents. Of course, leaders are not infinitely lived and are subject to the same actuarial risks that we all face. We explicitly model mortality, with each leader surviving each round of the game with probability λ and dying with probability $(1-\lambda)$. Leader mortality shapes relations between states. When sanctions are leader specific, leader turnover rejuvenates relations between states. Although the risk that the game ends is often given as an interpretation for discounting, we explicitly separate discounting from survival. All players have a common discount rate, δ. Unless interactions are between genuine unitary actors conflating discount rates and mortality mis-characterizes interstate relations. Although mortality means the end of a payoff stream for a leader, for citizens a leader's death can reinvigorate relations and improve their payoffs.

The compliance/punishment relationship is based on the credibility with which nation A can punish nation B. Nation A can enforce compliance by immediately punishing nation B if nation B is caught cheating, $y_{Bt} > h$. Although scholars have examined such mechanisms (see Downs and Rocke 1995; Green and Porter 1984; McGillivray and Smith 2004; Porter 1983), few interstate relations are based on immediate punishment. More usually when alleged violations occur nations engage in dialogue and allow each other the opportunity to make redress before punishments are applied. Institutions, such as the WTO's dispute resolution mechanism (www.wto.org), provide nations the opportunity to explicitly 'apologize.'

In international trade it is often difficult to distinguish between a policy designed to uphold a non-discriminatory domestic environmental standard, something allowed under WTO rules, and a discriminatory protectionist measure designed to inhibit trade, which is not allowed under WTO rules. The world is a noisy place. In terms of our model, the dispute resolution mechanism's judicial panel decides whether the observed policy (y_{Bt}) breaks the rules. That is the panel judges whether $y_{Bt} > h$. Once found in breach of an agreement, nation B has an opportunity to redress the harm caused to nation A. If nation B does not do so then the WTO authorizes nation A to take retaliatory sanctions. Rosendorff (2004) and Rosendorff and Milner (2001) show the ability to 'opt out' of the agreement and then apologize increases the robustness of the WTO's institutions. Not only does the ability to apologize allow nations to recover from accidental default, but Rosendorff argues nations are more willing to enter agreements when they know they can temporarily opt out in response to unanticipated domestic demands for protection.

The opportunity to 'apologize' is a common feature within compliance/punishment relations. For instance, in 1997 the US Congress threaten to remove Mexico's certification as an ally in the war on drugs. This threat followed a variety of incidents which suggested wide spread corruption in those Mexican law enforcement agencies responsible for drug enforcement. Perhaps most noticable was the arrest of General Jesus Gutierrez Rebello, head of newly created National Institute to Combat Drugs, on corruption charges. It was alleged that he worked for the Juarez drug cartel. Following these revelations, and the subsequent US threat to decertify Mexico, Mexican President Ernesto Zedillo implemented wide ranging reforms of Mexico's anti-drug forces which succeeded in restoring Mexico's reputation as a partner in the war on grugs (Hinojosa 2002).

We believe opportunities for compensation, rather than immediate punishment, are the norm and explicitly model these opportunities within our analysis of the compliance/punishment game mechanism. We say B has broken the rules if $y_{Bt} > h$. In the next period, nation B can 'apologize' to nation A through additional policy concessions, that is by setting $y_{Bt} \leq h_0$ where $h_0 < h$. For both technical convenience and to reflect increased monitoring of malfeasants, we assume that once nation B has been caught cheating ($y_{Bt} > h$) then its policies are perfectly monitored until its reputation is restored, i.e. $y_{Bt} = x_{Bt}$. In terms of the McCubbins and Schwartz's distinction (1987), this is to say that after the 'fire alarm' sounds 'police patrols' perfectly monitor behavior.

Compliance, punishment and credibility

Constructivist scholars, such as Chayes and Chayes (1995), Hathaway (2002, 2004) and Raustiala and Slaughter (2002), argue that as nations honor international agreements then the agreements develop legitimacy that help ensure future compliance. Proponents of these arguments are quick to cite the high levels of compliance within international organizations as evidence. Downs, Rocke and Barsoom 1996 (see also Gilligan 2004) argue instead that compliance is high because agreements are swallow, with agreements written to codify the policies that nations wish to implement anyway. In the context of our model, if the threshold is set at $h = 1$ then nation B can implement its preferred policy, never be found in violation of the agreement and so compliance and punishment are moot points. As an empirical matter, international agreements vary enormously, not only in the concessions they seek from parties but also in the their arrangements for monitoring (that is observing y_{Bt}) and specifying what actions are sanctionable in the event of a rules violation. The WTO specifies rules, a monitoring mechanism and enforcement through sanctions. Arms controls treaties, by contrast, although providing detailed rules and verification mechanisms, rarely if ever, specify what to do in the event of a violation. Other agreements, such as the Kyoto environmental agreements specify nothing beyond broad policy goals. That is Kyoto specifies neither a monitoring mechanism nor a punishment scheme in event that nations fail to comply. The result was been widespread non-compliance.

The principal problem in compliance/punishment arguments is that without a credible threat of punishment for non-compliance, nation B will never shift its policy position to comply with the rules. Yet, if it is costly to punish, as in our model ($c_A > 0$), then nation A has no myopic incentive to sanction nation B for non-compliance. To ensure compliance, not only does nation A need a stick with which to punish nation B, but nation A must also be able to commit to use the stick if B breaks the rules. A credible commitment to punish nation B is essential to A's ability to compel nation B.

Scholars have formally modeled the credibility problem from the context of the chain store paradox (Selten 1978). The original chain store model examines the behavior of a firm which is a monopolist in a number of local markets. Rival firms consider entry into each of these markets. If the competitor enters the market then the chain store can either compete at market prices with its rival or it can enter a price war with its new competitor by selling products at below cost. This price war imposes losses on both itself and the competitor in the local market. Of course, in any single local market it is not credible for the chain store to enter a price war. However if there are multiple markets and the chain store demonstrates its willingness to punish any rival who enters one of its markets then the chain store can deter rivals from entering in other local markets. Unfortunately, such a commitment lacks credibility. If rivals have already penetrated all but one of the chain store's local markets, then when a competitor enters the last market a price war has no impact on future entry decisions and only harms the chain store. Hence the chain store does not defend the last market. Now consider the penultimate local market. Whatever happens in this market, tomorrow the chain store will not defend the last market. Knowing this there is nothing to be gained from defending the penultimate market. Repeating this argument in each preceding market shows how the credibility of the price war threat unravels. Scholars have solved the credibility problem by assuming that a 'crazy' type exists who prefers a self defeating price war to competition (Kreps and Wilson 1982). However, it is troubling that the maintenance of A's reputation requires such a technical fix.

Alt, Calvert and Humes (1988) model hegemonic control by assuming that while it is generally costly to punish dissent, sometimes the hegemon can do so costlessly (for a discussion hegemony see Gilpin 1981; Keohane 1984 and Kindleberger 1986). By punishing any first round defector the hegemon might hope to convince states contemplating defection in the second round that it is a type of hegemon that is likely to be able to punish with low cost. Although Alt, Calvert and Humes shows that this mechanism can sometimes work, it is not credible for the hegemon to always punish in the first round and the maintenance of a reputation is stochastic.

Despite the reservations discussed above, the maintenance of a reputation for external audience deserves serious consideration. Therefore, we assume that nation A receives a payoff of R if it maintains an honest reputation. By an honest reputation we mean that nation A never allows nation B to go unpunished if nation B violates the rules and fails to apologize. If we think of an agreement as a commitment by N states to make concessions to nation A, then the payoff R can be thought of as $(N-1)$ times the payoff from having state B comply with A's rules. The payoff R is A's reward for a good external reputation. Although the payoff R increases the incentives of state A to maintain its reputation, we show that LSP allows the maintenance of credibility without reference to this exogenous payoff.

A MODEL OF COMPLIANCE AND PUNISHMENT.

In order to enforce compliance, nation A must commit to sanction B for past transgressions. To describe the conditionality of prior play of the game on current strategies we use Markovian state variables to describe the pertinent features of the history of previous play and we characterize Markov Perfect Equilibria (see Fundenberg and Tirole 1991 chapt. 13 for definitions). The Markov assumption means players base their strategies only on the state variable, which in this case contains the relevant information about past play. The behaviors we describe are sustainable as Subgame

Perfect Equilibria.

The formal model is an infinitely version of the compliance/punishment game:

Compliance/punishment game

1) β, the leader of nation B, picks policy $x_{B,t} \in [0,1]$. The outcome of the policy, $y_{B,t} = x_{B,t} + \varepsilon_{B,t}$, is observed by all players, where $\varepsilon_{B,t}$ is a stochastic error with distribution $F()$ if the state variable $Z_{Bt} = 0$, and $\varepsilon_{B,t} = 0$ otherwise.

2) α the leader of nation A decides whether or not to sanction B.

3) The citizens in nation A decide whether or not to replace their leader, α. The citizens in nation B then decide whether or not to keep leader β.[2]

4) Players receive payoffs based upon the outcome of the game.

5) Both leaders face mortality risks: with probability λ_j leader j survives; with probability $(1 - \lambda_j)$ leader j dies and is replaced by another leader.

6) The state variable $Z_t = (Z_{At}, Z_{Bt})$ is updated according to the passage of play.

The payoffs for the stage game are as follows: The citizens B and the leader β of nation B receive the payoff $U_B(x_t, s_t) = x_{B,t} - s_t c_B$, where x_{Bt} is leader β's policy choice, s_t is A's choice of whether to sanction and c_B is the cost of the sanctions if they are applied. Players in nation A receive the payoff $U_A(x, s) = (1 - x_{Bt}) - s_t c_A + R(1 - Z_{At})$ where R is the value of having a good reputation at the start of the period and c_A is the cost of sanctioning.

In addition to these payoff from the interaction between nations A and B, players receive domestic payoffs. Each leader, β and α, receives a payoff of Ψ if they are retained. The payoff Ψ represents the inherent value of holding office, which is assumed to be large: $\Psi >> 1$. Office holding is the primary motivation of leaders. The citizens in nation A pay a cost r_A to replace leader α. Similarly the citizens in nation B pay the cost r_B to replace β. $L_t = (L_{B,t}, L_{A,t})$ indicates which leaders changed at

[2] We assume there is an infinite pool of potential challengers such that, once removed, leaders have no prospects of returning. Challengers are identical to incumbents in every way.

time t.

Obviously the past histories of this game can be extremely complex. In subgame perfect equilibrium strategies can depend upon the entire history of play. While it is often easy to informally describe the pattern of play, a formal description of a player's strategy is often tedious and difficult to write down. Therefore, we restrict attention to only specific features of previous play that we describe with the state variables.

Z_{At} and Z_{Bt} are the standings of states A and B at time t. Initially $Z_{B0} = 0$, and $Z_{A0} = 0$, this is to say initially each nation is in good standing. Before precisely defining the evolution of the state variables, we describe them intuitively. In doing so we will use such terms as 'punishment' and 'caught cheating' which are intuitively clear features of equilibrium play.

Each nation starts in good standing and this good standing is restored with the replacement of its leader (either by the citizens or mortality). Strictly speaking a nation's standing belongs to its leader. Leader replacement reinvigorates a nation's integrity. Leader α (that is nation A) maintains a good standing unless she fails to sanction when she should. Once her standing is tarnished,$Z_{At} = 1$, her nation has a poor standing until she is replaced, .

Nation B's leader, β, also starts in good standing. If her observed policy actions are above the threshold, $y_{Bt} > h$, then she is said to have been caught cheating and her reputation becomes $Z_{Bt+1} = 1$. If $Z_{Bt} > 0$ and leader β apologizes, that is to say she plays $y_{Bt} \leq h_0$, then her reputation improves $Z_{Bt+1} = Z_{Bt} - 1$. If however she fails to apologize then the state variable increases: $Z_{Bt+1} = Z_{Bt} + 1$. Each time leader β fails to apologize her state variable increases and each time she apologizes it decreases. Nation B's reputation follows this pattern unless either nation A falls into poor standing or leader β is replaced in which case nation B's standing is refreshed: $Z_{Bt} = 0$.

All the equilibria we describe provide β with an opportunity to apologize before sanctioning starts. Downs and Rocke (1995), Green and Porter (1984), and Porter (1983) characterize equilibrium with punishment schemes that do not allow for apology. As with most infinitely repeated game, there are many equilibria. It is therefore important to argue why we focus on the equilibria that we do. As discussed above the apology feature is an important part of many real world interactions between states that helps provide flexibility to the institutional framework. Within this genre of equilibria we focus on maximizing enforcement and compliance of the rules. Of course from an economist's perspective compliance is not as important as efficiency. Green and Porter (1984) and Porter (1983), for example, characterize efficiency when there are underlying returns from nations A and B cooperating. However, the conflict in our model is inherently distributional and as such efficiency is of less concern. Indeed, in the compliance game Pareto efficiency implies any outcome without sanctions. Given our political focus on distributional concerns, we characterize how leader specific punishments and domestic institutions increase the parameter shape in which full compliance can be attained.

We now briefly describe play along the equilibrium path for the base case. Given she is in good standing ($Z_{Bt} = 0$) leader β plays $x_{Bt} = x^*$. If in the previous period she was caught cheating then she plays $x_{Bt} = h_0 < h$. In higher states she either apologizes ($x_{Bt} = h_0$ if $Z_{Bt} \leq \overline{Z}$) or sets $x_{Bt} = 1$ (if $Z_{Bt} > \overline{Z}$). At the point $Z_{Bt} > \overline{Z}$ leader β would have to apologize too many times to make restoring relations with A worthwhile.

Leader α sanctions if she is in good standing ($Z_{At} = 0$) and if nation B was caught cheating in the past and failed to apologize ($Z_{Bt} > 1$ or ($Z_{Bt} = 1$ and $y_{Bt} > h_0$)); otherwise leader α does not sanction. In the base case, that we examine first, citizens never replace their leaders.

Below we formally define the evolution of the state variable. Let event J describe the condition where A is expected to sanction but leader α fails to do so: $J =$((either

$Z_{Bt} > 1$ or $(Z_{Bt} = 1$ and $y_{Bt} > h_0))$ and $s_t = 0$ and $L_{At} = 0)$.

$$Z_{Bt+1} = \begin{cases} 0 \quad if & \begin{array}{c} Z_{At} = 1 \\ \text{or } L_{B,t} = 1 \\ \text{or } (Z_{Bt} = 0 \text{ and } y_{B,t} \leq h) \\ \text{or } J \end{array} \\ 1 \quad if & L_{B,t} = 0, Z_{At} = 0, Z_{Bt} = 0 \text{ and } y_{B,t} > h \text{ and not J} \\ Z_{Bt} + 1 \quad if & L_{B,t} = 0, Z_{At} = 0, Z_{Bt} > 0 \text{ and } y_{B,t} > h_0 \text{ and not J} \\ Z_{Bt} - 1 \quad if & L_{B,t} = 0, Z_{At} = 0, Z_{Bt} > 0 \text{ and } y_{B,t} \leq h_0 \text{ and not J} \end{cases}$$

$$Z_{At+1} = \begin{cases} 1 \quad if & \begin{array}{c} (Z_{At} = 1 \text{ and } L_{A,t} = 0) \text{ or} \\ (Z_{Bt} = 1, \, y_{B,t} > h_0, \, s_t = 0 \text{ and } L_{At} = 0) \text{ or} \\ (Z_{Bt} > 1 \, , \, s_t = 0 \text{ and } L_{At} = 0) \end{array} \\ 0 \quad if & otherwise \end{cases}$$

Strategies.—

We examine Markov strategies. Leader β chooses $x_t : Z_t \to [0, 1]$ where Z_t is Cartesian product $Z_{At} \times Z_{Bt}$. Leader α's decision to sanction is $s_t : Z_t \times Y_t \to \{0, 1\}$ where Y_t is β's observed policy choice. The citizens in nations A and B remove leader α and β respectively depending upon the state variables, the observed policy outcome and A's decision whether to sanction: $\gamma_a : Z_t \times Y_{B,t} \times S_t \to \{0, 1\}$ and $\gamma_B : Z_t \times Y_{B,t} \times S_t \times L_{At} \to \{0, 1\}$, where $S_t = \{0, 1\}$ is α's decision to sanction and $L_{At} = \{0, 1\}$ is leader replacement in state A.

Continuation values describe the expected payoff from playing the game under some strategy profile given the state variables. We use the notation $V_\alpha(Z_{At}, Z_{Bt})$ to described the expected payoff for leader α from playing the game starting in state (Z_{At}, Z_{Bt}). We use parallel notation for other actors. Given the principle of dynamic optimality, MPE require that there are no single period deviations from the prescribed strategy profile that are welfare improving for any player (Chiang 2000).

CREDIBILITY, COMPLIANCE AND PUNISHMENT.

LSP influence the credibility and depth of international agreements. We start with a base case in which citizens never remove their leader. Through exploration of key aspects of the equilibrium, we explain how LSP shape the credibility of punishments within the equilibrium. Formal proofs are contained in the appendix. For convenience we will frequently use a numerical example where $F(q) = 1 - e^{-\frac{q}{.05}}$, $\delta = .9$, $\lambda_B = .8$, $\lambda_A = .8$, $h_0 = 0$, and $h = .5$.

Base Case LSP (with no endogenous leader replacement)

Proposition 1: Base Case Leader Specific Punishments (equilibrium 1).

If the costs of leader removal are high ($r_A \geq V_A(0,0)\frac{1-\delta}{1-\delta+\delta\lambda_A}\delta$ and $r_B \geq \lambda_B\delta(V_B(0,0) - V_B(0,Z_{Bt}))$ for all Z_{Bt}), the cost of imposing sanctions for nation A are modest ($c_A \leq \delta\lambda_A\frac{(\delta\lambda_A(1-\delta\lambda_A)(1-\lambda_B))V_\alpha(0,0)-\delta\lambda_A(1-\lambda_B)\Psi+(1-\delta\lambda_A)R}{(1-\delta\lambda_A)(1+\delta\lambda_A(1-\lambda_B))}$), and sanctions impose high costs on nation B ($c_B \geq \frac{(-\lambda_B\delta p-1+\lambda_B\delta)h_0+\lambda_B\delta p+1-\delta\lambda_B x^*}{\lambda_B\delta p+1}$) then there exists a Markov Perfect Equilibrium where

$$1)\ x_{B,t} = \begin{cases} 1 & if \quad Z_{At} \neq 0 \text{ or } (Z_{At} = 0, Z_{Bt} > \overline{Z} \geq 1) \\ h_0 & if \quad Z_{At} = 0, \quad 0 < Z_{Bt} \leq \overline{Z} \\ x^* & if \quad Z_{At} = 0, Z_{Bt} = 0 \end{cases}$$

$$2)\ s_t = \begin{cases} 1 & if \quad Z_{At} = 0, \text{ and (either } Z_{Bt} > 1 \text{ or } (Z_{Bt} = 1 \text{ and } y_{B,t} > h_0)) \\ 0 & \text{otherwise} \end{cases}$$

3) $\gamma_A = 0$

4) $\gamma_B = 0$

and x^* solves $p\lambda_B\delta + 1 + \lambda_B\delta\frac{dp}{dx}(h_0 - x^*) = 0$, $p = 1 - F(h - x^*)$, $\frac{dp}{dx} = F'(h - x^*)$.

The statement of the equilibrium also requires the following definitions: $\overline{Z} \geq 1$ is the largest Z_{Bt} such that $V_\beta(0, Z_{Bt}) \leq \frac{1-c_B+\Psi}{1-\delta\lambda_B}$, where $V_\beta(0,1) = h_0+\Psi+\lambda_B\delta V_\beta(0,0)$, and if $Z_{Bt} = Z \geq 2$ then $V_\beta(0,Z) = \delta^Z\lambda^Z V_\beta(0,0) + h_0\delta^{Z-1}\lambda^{Z-1} + \sum_{\tau=0}^{Z-2}\lambda^\tau\delta^\tau(h_0 - c_B)$

$+ \Psi \frac{1-\delta^Z\lambda^Z}{1-\delta\lambda_B}$ *and* $V_\beta(0,0) = \frac{x^*+\Psi+\lambda_B\delta p(h_0+\Psi)}{(1-\lambda_B\delta)(p\lambda_B\delta+1)}$. $V_A(0,0) = \frac{1-x^*+\lambda_B\delta p(1-h_0)+R(1+\delta\lambda_B p)}{(1-\delta)(\lambda_B\delta p+1)}$. $V_B(0,0) = \frac{x^*+h_0\lambda_B\delta p}{(1-\delta)(\lambda_B\delta p+1)}$, $V_B(0,1) = h_0 + \delta V_B(0,0)$ *and if* $1 < Z_{Bt} = Z \leq \overline{Z}$ *then* $V_B(0,Z) = V_B(0,0)\delta(\lambda_B^{Z-1}\delta^{Z-1} + \sum_{j=1}^{Z-1}(1-\lambda_B)\lambda_B^{j-1}\delta^{j-1}) + h_0\sum_{j=1}^{Z}\delta^{j-1}\lambda_B^{j-1} - c_B\sum_{j=1}^{Z-1}\delta^{j-1}\lambda_B^{j-1}$ *and if* $Z_{Bt} > \overline{Z}$ *then* $V_B(0,Z_{Bt}) = \frac{1-c_B+\delta(1-\lambda_B)V_B(0,0)}{1-\delta\lambda_B}$.

Proof in appendix.

In the base case, citizens never remove their leaders: $\gamma_A = 0$ and $\gamma_B = 0$. Leader α sanctions if and only if nation B has crossed the line ($y_{Bt} > h$) and not apologized. In the initial state leader β plays policy $x_{Bt} = x^*$; if she crosses the line ($y_{Bt} > h$) then she apologizes in the next period ($y_{Bt} = h_0$). Should A ever lose its good standing ($Z_{At} = 1$) then β maximizes policy, $x = 1$. Should β need to apologize too many times, such that restoring good relations is not worthwhile, ($Z_{At} = 0$, $Z_{Bt} > \overline{Z}$) then β maximizes policy ($x = 1$) and accepts sanctions will be applied.

The base case equilibrium has several similarities to the chain store paradox. The maintenance of the system of compliance and punishment relies on the credibility that A continues to sanction even in states $Z_{Bt} > \overline{Z}$; that is leader α continues to sanction even when nation B's leader never apologizes and never complies with the rules in these Markov states. If leader α cannot commit to sanction under every possible circumstance then the whole system of punishments unravels. To illustrate this logic, suppose that β will never apologize to restore good relations once $Z_{Bt} > 10$. Once β has failed to comply sufficiently that she will never in the future apologize, $Z_{Bt} > 10$, then leader α has no prospect of obtaining future concessions from β and sanctions serve no purpose. Suppose that, unlike the case in equilibrium 1, leader α does not sanction in these pointless states. Next consider behavior in state $Z_{Bt} = 10$. Leader β can restore relations by apologizing 10 times whilst being sanctioned and then set policy to x^* in future rounds. Alternatively, leader β can fail to apologize one more time such that $Z_{Bt+1} > 10$. Under in this latter condition α will never sanction again despite β picking B's most desired policy, $x = 1$. Since, from β's perspective, the latter is more desirable than the former, β will break the rules this final time. Leader α finds it useless to sanction in state 10 since the threat does not influence β's behavior in state 10. Therefore, if leader α will not sanction in states $Z_{Bt} > 10$, then α should not sanction in state 10. Using an inductive argument this process can be repeated: if α does not sanction in state Z then β does not apologize in state $Z-1$ and neither does α sanction. If leader α can not commit to sanction in *all states* then the whole system of threats unravels. Ensuring obedience to the rules requires credible punishment.

The continuation value $V_\beta(Z_{At}, Z_{Bt})$ is the expected value for leader β of playing the game starting in state (Z_{At}, Z_{Bt}). In the initial state, ($Z_{At} = 0, Z_{Bt} = 0$), leader β plays policy x^*. With probability $p = \Pr(x^* + \varepsilon_{it} > h) = 1 - F(h - x^*)$ nation B is observed to cross the line and breach the threshold h. That is, nation B is observed to break the rules. Following such a breach, in the next period the state variable will be ($Z_{At} = 0$, $Z_{Bt} = 1$) and nation B apologizes by setting $x_{Bt} = h_0$. This restores the state to the initial ($Z_{At} = 0$, $Z_{Bt} = 0$). Leader β's continuation value in state $(0, 1)$ is $V_\beta(0, 1) = \Psi + h_0 + \delta\lambda_B V_\beta(0, 0)$, which is the value of office holding and the value of the policy h_0 plus the discounted value of playing the game tomorrow given that the leader survives, which occurs with probability λ_B.

Leader β's continuation value for the initial state (0,0) is $V_\beta(0, 0) = (\Psi + x^*) + p\delta\lambda_B V_\beta(0, 1) + (1-p)\delta\lambda_B V_\beta(0, 0)$, that is the value of office holding and policy x^* plus the discounted (and mortality risk decreased) payoff associated with future play, which with probability p is $V_\beta(0, 1)$ and with probability $(1 - p)$ is $V_\beta(0, 0)$. Substituting in the value for $V_\beta(0, 1)$ yields that $V_\beta(0, 0) = \frac{\Psi + x^* + p\delta\lambda_B(\Psi + h_0)}{(1-\delta\lambda_B)(p\delta\lambda_B+1)}$, where $p = 1 - F(h - x^*)$. To ensure it maximizes her payoff, leader β picks the value of x^* that maximizes $V_\beta(0, 0)$, which implies the first order condition $p\lambda_B\delta + 1 + \lambda_B\delta\frac{dp}{dx}(h_0 - x) = 0$. For our numerical example $x^* = 0.417$ and the probability that β is caught cheating is $p = 0.189$.

Leader β chooses to apologize if she is caught cheating ($Z_{Bt} = 1$). Alternatively, she might delay the restoration of cooperation by maximizing policy, or perhaps never apologize. However, when the cost of sanctions, c_B, is sufficiently high, specifically $c_B \geq \frac{(-\lambda_B\delta p - 1 + \lambda_B\delta)h_0 + \lambda_B\delta p + 1 - \delta\lambda_B x^*}{\lambda_B\delta p + 1} = 0.736$, then β prefers to apologize.

In this base case equilibrium, leader β apologizes because she knows if she does not then she will be subject to indefinite sanction by nation A. Yet, the whole equilibrium is supported on the threat that nation A will punish non-compliance in every state. How credible is this threat? In order for A to maintain its credibility requires that it is prepared to sanction defectors. Unfortunately, when the defector is prepared to endure sanctions indefinitely rather than apologize then A gains nothing from sanctions which have no possibility of success. A's credibility relies upon being able to commit to sanction. Only if the cost of sanctioning is sufficiently low can α credibly commit. Specifically, if $c_A \leq \frac{(1-\lambda_B\delta)(\lambda_B\delta p+1)R + \delta\lambda_A(1-\lambda_B)x^* + \delta^2\lambda_A\lambda_B p(1-\lambda_B)h_0}{(1-\lambda_B\delta)(\lambda_B\delta p+1)} = 0.953R + 0.202$.

To understand α's credibility it is essential to consider the origins of this limit on the cost of sanctioning. If α fails to sanction when it should punish nation B ($Z_{Bt} > 1$ or ($Z_{Bt} = 1$ and $y_{Bt} > h_0$)) then α loses all her credibility. In future rounds no nation will be accommodating to nation A. Hence α's payoff is $1 - x_{Bt} + \Psi + R + \delta\lambda_A(V_\alpha(1,0))$, where $V_\alpha(1,0) = \frac{\Psi}{1-\delta\lambda_A}$. If α sanctions then she retains her credibility and her payoff is $1 - x_{Bt} - c_A + \Psi + R + \delta\lambda_A(V_\alpha(0, Z_{Bt+1}))$, where Z_{Bt+1} describes the state in following rounds given play in the current round. The worst this payoff can be is in a state $Z_{Bt+1} > \overline{Z}$, at which point nation B has to apologize too many times to make apology worthwhile. Instead in these states, β sets $x_{Bt} = 1$ and endures sanctions in future periods. In such a setting, leader α continues sanctioning until leader β dies and relations are restored. Hence only if $c_A \leq \delta\lambda_A(V_\alpha(0, Z_{Bt}) - V_\alpha(1,0))$ for all Z_{Bt} is leader α's commitment to sanction credible.

In the base case, credibility is obtained through two paths. First, A does not have to sanction indefinitely because leader β is mortal and, if she dies, or is otherwise replaced, then sanctions end as nation B's record as an unapologetic non-complier is wiped clean. This reduces the expected length of sanctions. Second, if nation A does not sanction defector B then it loses all credibility vis-à-vis other states. As a result, nation A loses the payoff R, which it receives for maintaining its external reputation. This payoff of R greatly enhances the maintenance of credibility. However, as $R \to 0$ and leader β becomes more stable in office, $\lambda_B \to 1$, then sanctions need to become costless in order for nation A to credibly maintain its commitment to sanction unapologetic transgressors.

The base case suffers from the same credibility problem as the chain-store. While the equilibrium demonstrates that A is willing to sanction nation B's non-compliance to keep other states in compliance, if all states were in non-compliance then A would have nothing to gain from sanctioning nation B. But then since A has no credible incentive to sanction the last nation, it could not credibly commit to punish the penultimate nation, and so A's credibility unravels is a manner akin to that described by the chainstore paradox. In the base case equilibrium, nation A's ability to discipline the system relies upon it never having to discipline many states at a time. While it may well be true that A's interactions signal how other states can expect to interact with A, it creates a house of cards because A's credibility is, in effect, created by reference to an exogenous payoff (R) and is not internally derived. However, if leader α is easily replaced then leader specific punishments resolve nation A's commitment problem.

Next we examine how linking a leader's survival in office to her willingness to sanction through leader specific punishment strategies enables domestically accountable leaders to more credibly commit to punishment non-compliance.

Domestically Accountable Leaders and Credible Punishment

If leader α ever fails to punish a non-apology then she loses her integrity and it becomes impossible for her to credibly commit to sanction in the future. Nation B will never again comply with the rules once α has revealed her inability to commit to punish. As long as α remains the leader of nation A, nation B will not comply. Yet under leader specific punishment strategies, nation A's reputation is restored by replacing leader α. By deposing their leader, the citizens in nation A restore their national reputation and restart nation B's compliance. If the value of restoring their nation's reputation, which is derived from B's compliance, is greater than the cost of leader removal then the citizens depose their leader if she fails to punish non-compliance appropriately.

If the costs of leader removal in nation A are sufficiently modest, then leader α can commit to punish nation B's non-compliance and failure to apologize. If leader α does not sanction, then she will be domestically deposed. Since retaining office is the primary goal of political leaders and a failure to punish appropriately costs a leader her job, leaders can commit to punish. We examine this commitment through Equilibrium 2. Since in most regards this equilibrium 2 is identical to the base case (equilibrium 1), we state the differences between equilibria 1 and 2 below and defer the formal characterization to the appendix.

Equilibrium 2: If the cost of removing leader α is low ($r_A \leq \delta\lambda_A \frac{1-\delta}{1-\lambda_A\delta}\frac{1-x^*+\lambda_B\delta p(1-h_0)+R(1+\delta\lambda_B p)}{(1-\delta)(\lambda_B\delta p+1)} = 1.628\,3{+}2.571R$), and $c_A \leq \Psi+\lambda_A\delta\lambda_B R+ \delta\lambda_A(1-\lambda_B)\frac{(1+\delta\lambda_A p\lambda_B)(\Psi+R+1)-\delta\lambda_A p\lambda_B h_0-x^*}{(1-\delta\lambda_A)(1+\delta\lambda_A p\lambda_B)} = 1.514\Psi+1.090R+0.321$ then there exists a MPE in which leader α always sanctions states caught cheating that fail to apologize (either $Z_{Bt}=1$ and $y_{Bt}>h$ or $Z_{Bt}>1$) and the citizens in nation A remove α should she fail to sanction when she should.

The major difference between equilibrium 2 and its base case predeccessor (equilibrium 1) is that if leader α fails to sanction when B does not apologize then the citizens remove her. Since leaders are primarily driven by office holding motives, they can be relied upon to sanction even when it is costly to do so. As nation A's institutions become more inclusive, and hence leader removal becomes easier, A's threats to sanction become more credible and adherence to the rules becomes easier to maintain. Indeed if one believes, as we do, that the dominant goal of all political leaders is to retain office, leader α will readily sanction under nearly all conditions. Once leaders are domestically accountable and removed for not appropriately sanctioning then sanctions become fully credible. Unlike the chain store paradox or the base case equilibrium 1, there is little danger of equilibrium 2 "unravelling." Leaders sanction to preserve their jobs, not because the sanctions are an effective foreign policy.

In the international relations literature, domestic costs, such as the loss of office, paid by leaders who renege on promises are referred to as audience costs (Fearon 1994; Guisinger and Smith 2002; Leeds 1999; Smith 1998; Schultz 2002). While much of the literature considers these costs as exogenous, here the costs are endogenously derived with the citizens having incentives to remove leaders who do not follow through on sanctions. Indeed the leader specific strategies encourage citizen to punish leaders who fail to sanction, even when the sanctions are not in the citizens interests. There is a range of sanction costs, specifically $\delta\lambda_A \frac{(1-\delta\lambda_A\lambda_B)(p\delta\lambda_A\lambda_B+1)R+\lambda_A\delta(1-\lambda_B)(1-x+p\delta\lambda_A\lambda_B(1-h_0))}{(p\delta\lambda_A\lambda_B+1)(1-\delta\lambda_A\lambda_B+\delta\lambda_A)(1-\delta\lambda_A)}$ $< c_A \leq \Psi + \lambda_A\delta\lambda_B R + \delta\lambda_A(1-\lambda_B)\frac{(1+\delta\lambda_A p\lambda_B)(\Psi+R+1)-\delta\lambda_A p\lambda_B h_0 - x^*}{(1-\delta\lambda_A)(1+\delta\lambda_A p\lambda_B)}$ (which for our numerical example translates into $0.953\,05R+0.201\,95 < c_A \leq 1.514\Psi+1.090R+0.321$), that a unitary actor nation (or the citizens of that state acting for themselves) would be unwilling to pay to maintain their honest reputation, (i.e. the cost c_A is too high to support equilibrium 1), but for which equilibrium 2 exists. That is to say, LSP allow leaders to commit themselves to a much greater extent than could either a unitary actor or the citizens. Of course, to some extent this means that accountable leaders act contrary to the interests of their citizens by punishing even when the cost of doing so is high. However, precisely because accountable leaders can so credibly threaten to punish, they induce high levels of compliance and so rarely have to carry out these threats. In this context, the persistence of the Clinton and Bush administrations' maintenance of economic sanctions against Iraq despite their enormous humanitarian costs and their obvious failure to hurt Saddam Hussein's regime becomes easier to explain (Rieff 2003).

Democracy, or other institutional arrangements that make leader removal easy, increase nation A's credibility and so enable A to induce greater compliance.[3]

The Compliance of Accountable Leaders and LSP

When leader α is politically accountable then the credibility of nation A's threat to punish transgressors is enhanced which makes compliance easier to maintain. Democratic states such as the US should be better able to commit to punish their errant satellites than authoritarian states. Next we consider the consequences of political institutions in nation B. If nation B's leader is relatively accountable, as would be the case if B were democratic, then leader β is more likely to apologize for any transgressions and is potentially less likely to be caught cheating in the first place. We demonstrate these phenomena through an examination of two equilibria. In the first (equilibrium 3) the cost of leader removal in state B is moderate and citizens give their leader β the opportunity to apologize for having been caught cheating. If, however, leader β fails to apologize then her citizens depose her. In the second (equilibrium 4) the cost of leader removal in state B is sufficiently low that citizens remove any leader caught cheating even before she has had the opportunity to apologize. Since in this latter equilibrium, being caught cheating is equivalent to removal, leader β

[3] Although not presented here, LSP make sanctions credible even in the case where nation i would prefer to endure sanctions rather than appologize ($\overline{Z} = 0$). No equilibria of this type exist for our numerical example.

chooses equilibrium levels of x well below the threshold h to reduce the chance of being caught cheating.

Proposition 3: Leader Specific Punishments with Domestically Accountable Leaders and Single Apology (equilibrium 3).

If the cost of leader removal in nation A is low ($r_A \leq \delta\lambda_A \frac{1-\delta}{1-\lambda_A\delta} V_A(0,0)$), the cost of leader removal in nation B is moderate ($r_B \geq \delta\lambda_B((1-\delta)V_B(0,0) - h_0)$, $r_B \leq \delta\lambda_B \frac{-1+c_B+(1-\delta)V_B(0,0)}{1-\delta\lambda_B}$ and $r_B \leq -\delta\lambda_B(1+\delta\lambda_B)h_0 + \lambda_B\delta(1-\delta)(\delta\lambda_B+1)V_B(0,0)$), and $c_A \leq \Psi + \lambda_A\delta V_\alpha(0,0)$ and $c_A \leq \Psi + \lambda_A\delta V_\alpha(0,1)$ then there exists a Markov Perfect Equilibrium where

$$1)\ x_{Bt} = \begin{cases} 1 & if \quad Z_{At} \neq 0 \text{ or } (Z_{At}=0,\ Z_{Bt}>2) \\ h_0 & if \quad Z_{At}=0, \text{ and } (Z_{Bt}=1 \text{ or } Z_{Bt}=2) \\ x^* & if \quad Z_{At}=0,\ Z_{Bt}=0 \end{cases}$$

$$2)\ s_t = \begin{cases} 1 & if \quad Z_{At}=0, \text{ and (either } Z_{Bt}>1 \text{ or } (Z_{Bt}=1 \text{ and } y_{B,t}>h_0)) \\ 0 & \text{otherwise} \end{cases}$$

$$3)\ \gamma_A = \begin{cases} 1 & if \quad Z_{At}=1 \text{ or condition } J \\ 0 & \text{otherwise} \end{cases}$$

$$4)\ \gamma_B = \begin{cases} 1 & if \quad Z_{At}=0 \text{ and } (Z_{Bt}>2 \text{ or } (Z_{Bt}=1,2 \text{ and } y_{Bt}>h_0)) \text{ and } L_{At}=0 \\ 0 & \text{otherwise} \end{cases}$$

where x^* solves $p\lambda_B\delta + 1 + \lambda_B\delta\frac{dp}{dx}(h_0 - x^*) = 0$, $\frac{dp}{dx} = F'(h - x^*)$.

The statement of the equilibrium also requires the following definitions: $V_A(0,0) = \frac{\delta\lambda_B p(1-h_0)+\delta\lambda_B pR+1-x^*+R}{(1-\delta)(1+\delta\lambda_B p)}$. $V_\alpha(0,1) = (1-h_0)+R+\Psi+\delta\lambda_A(V_\alpha(0,0))$ and $V_\alpha(0,0) = \frac{(1+\lambda_A\delta\lambda_B p)(R+\Psi)+(1-x^*)+\lambda_A\delta\lambda_B p(1-h_0)}{(1-\lambda_A\delta)(1+\lambda_A\delta\lambda_B p)}$. $V_B(0,0) = \frac{x^*+h_0\delta\lambda_B p}{(1-\delta)(1+\delta\lambda_B p)}$.

Proof in the appendix.

On the equilibrium path, equilibrium 3 is identical to the base case. What differs is that this equilibrium exists for a far wider range of sanctions costs. As we have already shown above, when α is easily removed nation A can commit to sanction transgressors more credibly that would be the case otherwise. When nation B also has accountable political institutions then it can credibly commit to apologize under a wider range of conditions than would be the case if leader β were unaccountable.

If leader β fails to apologize then nation A sanctions nation B. However, given the leader specific nature of punishments, the citizens in nation B can avoid the consequences of sanctions by replacing leader β. If the citizens depose leader β then they pay cost r_B. The criteria that $r_B \geq \delta\lambda_B((1-\delta)V_B(0,0) - h_0)$ ensures that the citizens in nation B prefer to give their leader the opportunity to apologize rather than depose her immediately, providing the apology can be made before sanctions. The conditions $r_B \leq \delta\lambda_B \frac{-1+c_B+(1-\delta)V_B(0,0)}{1-\delta\lambda_B}$ ensures that citizens in nation B prefer to replace their leader rather than be subject to sanctions in subsequent periods. This later condition ensures that the citizens replace their leader if she refuses to apologize ($y_{Bt} > h_0$) after having previously been caught cheating.

The threat to leader β's tenure in office ensures she apologizes. If β apologizes having transgressed the threshold in the previous period ($Z_{Bt} = 1$), then her payoff is $h_0 + \Psi + \delta\lambda_B V_\beta(0,0)$. If she does not apologize then she is removed from office, which is worth $x_{Bt} - c_B$. When office holding motivations dominate, leader β can be guaranteed to apologize to save her job.

Accountable political institutions help ensure the survival of international regimes. Democracies and other forms of accountable government produce incentives for hegemonic leaders to be credible in their threats and for leaders in other states to 'play by the rules'. While autocrats might choose not to apologize when they are detected cheating, democrats rush to apologize. As Rosendorff and Milner (2001) show, democratic leaders prefer institutional arrangements that give them the opportunity to apologize. Leaders are happy to apologize rather than lose their jobs. The ability to apologize prior to punishment means leaders need not distort their policies in fear of any accidental error leading immediately to sanctions. However, this flexibility is lost if leader replacement in nation B is extremely easy as citizens replace leaders before they can apologize.

In equilibrium 3 the citizens in nation B remove leaders who fail to apologize in order to avoid sanctions. However, if the cost of leader removal is extremely low, then the citizens might not give their leader the opportunity to apologize by deposing her immediately. This results in leader β picking a lower level of x to reduce the probability of her being caught over the threshold ($y_{Bt} > h$).

Proposition 4: Leader Specific Punishments with Domestically Accountable Leaders and Immediate Removal (equilibrium 4).

If the cost of leader removal in nation A is low ($r_A \leq \delta\lambda_A \frac{1-\delta}{1-\lambda_A\delta} V_A(0,0)$)*, the cost of leader removal in nation B is low* ($r_B \leq \frac{\lambda_B\delta}{1-\lambda_B\delta}(-1 + c_B + (1-\delta)V_B(0,0))$ *and* $r_B \leq -\lambda_B\delta h_0 + \delta\lambda_B(1-\delta)V_B(0,0)$)*, and* $c_A \leq \Psi + \lambda_A\delta V_\alpha(0,0)$*, and* $c_B \geq 1 - h_0 - \Psi - \lambda_B\delta V_\beta(0,0)$ *then there exists a Markov Perfect Equilibrium where*

$$1)\ x_{B,t} = \begin{cases} 1 & if \quad Z_{A,t} \neq 0 \text{ or } Z_{Bt} > 1 \\ h_0 & if \quad Z_{A,t} = 0,\ Z_{Bt} = 1 \\ x^* & if \quad Z_{A,t} = 0,\ Z_{Bt} = 0 \end{cases}$$

$$2)\ s_t = \begin{cases} 1 \quad if \quad Z_{A,t} = 0 \text{ and (either } Z_{B,t} > 1 \text{ or } (Z_{B,t} = 1 \text{ and } y_{B,t} > h_0)) \\ 0 \qquad\qquad \text{otherwise} \end{cases}$$

$$3)\ \gamma_A = \begin{cases} 1 \quad if \quad Z_{A,t} = 1 \text{ or condition J} \\ 0 \qquad\qquad \text{otherwise} \end{cases}$$

$$4)\ \gamma_B = \begin{cases} 1 \quad if \quad Z_{At} = 0 \text{ and } (Z_{B,t} \geq 2 \text{ or } (Z_{Bt} = 0 \text{ and } y_{Bt} > h) \text{ or } (Z_{Bt} = 1 \text{ and } y_{Bt} > h_0)) \\ 0 \qquad\qquad \text{otherwise} \end{cases}$$

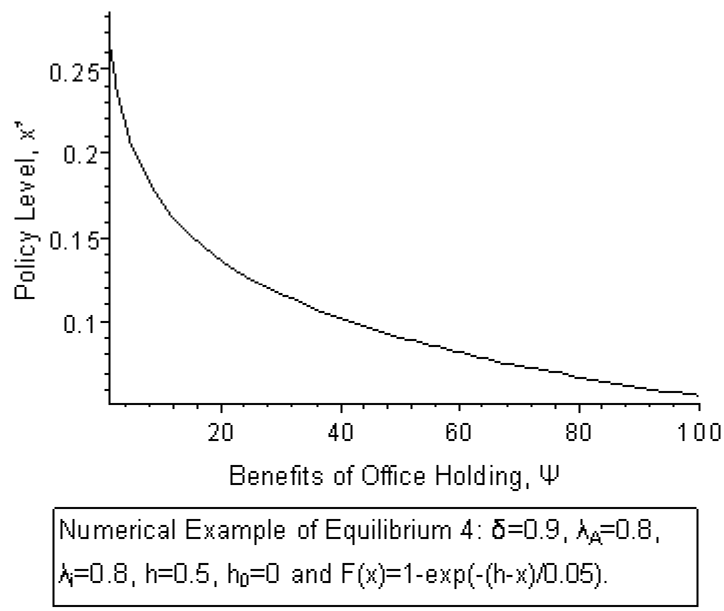
where x^* *solves* $(1 - \lambda_B\delta + \lambda_B\delta p) - (\Psi + \lambda_B\delta x)\frac{\partial p(x)}{\partial x} = 0$*,* $p = 1 - F(h - x^*)$ *and* $\frac{\partial p(x)}{\partial x} = F'(h - x)$*.*

The statement of the equilibrium also requires the following definitions:

$V_\alpha(0,0) = \frac{\Psi + R + (1-x^*)}{1-\delta\lambda_A}$*.* $V_\beta(0,0) = \frac{x^* + \Psi(1-p)}{1-\lambda_B\delta+\lambda_B\delta p}$*.* $V_A(0,0) = \frac{(1-x^*)+R}{1-\delta}$*.* $V_B(0,0) = \frac{x^* - r_B p}{1-\delta}$*.*

Proof in the appendix.

FIG. 1. Nation B's Policy Choice, x*, and the Value of Office Holding.



In equilibrium 4 the citizens of nation B remove their leader if she is ever caught cheating: $y_{Bt} > h$. By doing so they avoid having to apologize to nation A before restoring cooperation. This equilibrium requires that the cost of leader removal is low: $r_B \leq \frac{\lambda_B \delta}{1-\lambda_B \delta}(-1 + c_B + (1-\delta)V_B(0,0))$ and $r_B \leq -\lambda_B \delta h_0 + \delta \lambda_B (1-\delta) V_B(0,0)$.

Given that being caught cheating costs leader β her job, β sets policy to a lower value than in equilibria 1, 2 and 3 above. Figure 1 graphs x^* against the value of office holding for our numerical example.

If, for instance, $\Psi = 20$ and $R = 5$ then $x^* = 0.136\,62$ which means the probability of being observed in excess of the threshold is $p = .0007$. This equilibrium exists providing $c_A \leq \Psi + R + \lambda_A \delta \frac{\Psi + R + (1-x)}{1-\delta\lambda_A} = 91.\,506$, $r_A \leq \delta \lambda_A \frac{1-\delta}{1-\lambda_A \delta} \frac{(1-x)+R}{1-\delta} = 15.\,077$ and $r_B \leq \max\{0.09837 - 0.719 c_B, -2.\,220 + 0.001794 c_B\}$. Once leaders are politically accountable, compliance with international rules of conduct is easily maintained.

CONCLUSIONS

Credible punishments are essential for enforcing compliance. This paper explicitly analyzes how leader specific punishments, leader mortality and the ease of leader replacement affects the conditions under which compliance can be maintained through the threat of sanctions. The focus on individual leaders, rather than the nations they represent, shows that many credibility problems are solved through leader specific punishments. If α, the leader in nation A, directs punishment for non-compliance against state B only for as long as the aberrant leader of state B (β) remains in power, then A's threats of punishment are more credible that would be the case if A and B were the standard unitary actors present in traditional models of international relations. Since leader β is not immortal, the leader in state A knows that a commitment to leader specific punishment is not open ended commitment. Further, the desire to end sanctions encourages the citizens of state B to remove aberrant leaders. When domestic political institutions make such removal easy then non-compliance jeopardizes β's term in office. To avoid lose of tenure β is likely to be compliant and to apologize quickly for any accidental non-compliance. Inclusive domestic political institutions enhance compliance.

Domestic institutions that make leader removal easy also strengthen a leader's commitment to punish non-compliance. If leader α's failure to punish non-compliance jeopardizes her nation's future credibility (and hence ability to enforce future compliance), then the citizens of nation A are keen to replace α. For this reason, easily replaced leaders zealously maintain their integrity by punishing non-compliance, even if it is widely suspected that sanctions will not produce future compliance. Such commitments can unfortunately leads to poor ex post outcomes, with nation A maintaining damaging sanctions that have little prospect of successfully forcing apology or compliance from B. However, in expectation, these ex post loses are more than offset by the improvement in the ex ante probability that A will successfully compel nation B to comply with A's wishes in the first place.

The movement away from the unitary actor conception of nations toward a consideration of the survival induced motivations of political leaders is a growing trend in the study of international relations (see for example Bueno de Mesquita et al 2003 and Goemans 2000). While traditional research paradigms, such as realism and liberalism, have equated the interests of the leader and the state, their motivations differ. There is growing recognition of this distinction within international relations scholarship. As this paper shows, conditioning behavior on leaders, rather than the abstract concept of the state, increases the depth and breadth of compliance that can be maintained. This has important for implications for the design of international institutions and agreements. It also suggests that nations can greatly leverage the concessions they can extract from other states by explicitly focusing threats of punishments for non-compliance against leaders.

While there is some evidence of leader specific punishments in some policy areas (McGillivray and Stam 2004; McGillivray and Smith 2004), the use of leader specific policies is by no means universal. Although the analyses here suggest such policies enhance compliance, a valuable feature in the design of international agreements, given the prescriptive nature of our results, it is beholden on us discuss whether such mechanisms are really practical. By explicitly pointing to leader specific punishments, leaders, at least democratic ones, can dramatically increase what they can commit to. Yet, such improvements come with costs as leader specific punishments require leaders to stake their tenure in office on their ability to comply and punish non-compliance. If, as we believe, leaders are primarily driven by office holding goals, then leaders avoid any potential risk to tenure. This suggests that leaders are reluctant to stake their survival in office on their ability to honor an international agreement. Yet, if such a commitment improves the quality of citizens' lives in one nation, then it is attractive for potential challengers to make such offers. States with competitive political processes are the states that are most likely to adopt any form of leader specific punishment, and it is such nations that will obtain concessions from other states and be able to credibly commit to deep international agreements.

APPENDIX

In this section we formally characterize the equilibria discussed above and describe the conditions under which each of the equilibria exist. Since we discussed many of the pertinent feature of the proofs in the main text, the presentation of proofs take the following terse form. First we calculate the continuation values for each actor under each possible state. Second, for each actor we show that no single period deviations from the prescribed equilibrium path is utility improving.

Proof of Proposition 1: Base Case Leader Specific Punishments (equilibrium 1).—

We structure the proof as follows. First we calculate the continuation values for each player under each possible state given the equilibrium. Second, player by player we examine every possible one period deviation from the equilibrium path in each possible state and show that no such deviations are utility maximizing.

Continuation values of leader α. $V_\alpha(Z_{At}=1,.) = \Psi + \lambda_A \delta V_\alpha(Z_{At}=1,.) = \frac{\Psi}{1-\delta\lambda_A}$

$V_\alpha(0,0) = (1-x^*) + R + \Psi + \delta\lambda_A((1-\lambda_B p)V_\alpha(0,0) + p\lambda_B V_\alpha(0,1))$

$V_\alpha(0,1) = (1-h_0) + \Psi + R + \delta\lambda_A V_\alpha(0,0)$ where $p = \Pr(y_{Bt} = x^* + \varepsilon_{it} > h) = 1 - F(h-x^*)$

Substitution of $V_\alpha(0,1)$ yields $V_\alpha(0,0) = \frac{(1+\delta\lambda_A p\lambda_B)(\Psi+R+1)-\delta\lambda_A p\lambda_B h_0 - x^*}{(1-\delta\lambda_A)(1+\delta\lambda_A p\lambda_B)}$. If $2 \leq \overline{Z}$ then

$V_\alpha(0,2) = (1-h_0) - c_A + \Psi + R + \delta\lambda_A(\lambda_B V_\alpha(0,1) + (1-\lambda_B)V_\alpha(0,0)) = (1-h_0) - c_A + \Psi + R + \delta\lambda_A(\lambda_B((1-h_0) + \Psi + R + \delta\lambda_A V_\alpha(0,0)) + (1-\lambda_B)V_\alpha(0,0))$. Similarly, if $3 \leq \overline{Z}$ then $V_\alpha(0,3) = (1-h_0) - c_A + \Psi + R + \delta\lambda_A(\lambda_B((1-h_0) - c_A + \Psi + R + \delta\lambda_A(\lambda_B((1-h_0) + \Psi + R + \delta\lambda_A V_\alpha(0,0)) + (1-\lambda_B)V_\alpha(0,0))) + (1-\lambda_B)V_\alpha(0,0))$.

More generally if $1 < Z_{Bt} = Z \leq \overline{Z}$

$V_\alpha(0,Z) = -\sum_{j=1}^{Z-1}(\delta\lambda_A\lambda_B)^{j-1}c_A + (\Psi + R + 1 - h_0)\sum_{j=1}^{Z}(\delta\lambda_A\lambda_B)^{j-1} + (\sum_{j=1}^{Z-1}(1-\lambda_B)(\delta\lambda_A\lambda_B)^{j-1} + (\delta\lambda_A\lambda_B)^{Z-1})\delta\lambda_A V_\alpha(0,0)$.

if $Z_{Bt} > \overline{Z}$ then

$V_\alpha(0, Z_{Bt}) = (-c_A) + R + \Psi + \delta\lambda_A[\lambda_B V_\alpha(0, Z_{Bt}+1) + (1-\lambda_B)V_\alpha(0,0)]$ where $V_\alpha(0, Z_{Bt}+1) = V_\alpha(0, Z_{Bt})$, so for states $Z_{Bt} > \overline{Z}$, $V_\alpha(0, Z_{Bt}) = \frac{-c_A + R + \Psi + \delta\lambda_A(1-\lambda_B)V_\alpha(0,0)}{1-\delta\lambda_B\lambda_A}$.

Continuation values for leader β of nation B. If $Z_{At} = 1$ then $V_\beta(1,.) = 1 + \Psi + \delta\lambda_B[\lambda_A V_\beta(1,0) + (1-\lambda_A)V_\beta(0,0)]$

If $Z_{At} = 0$ and $Z_{Bt} = 0$ then

$V_\beta(0,0) = x^* + \Psi + \lambda_B\delta(pV_\beta(0,1) + (1-p)V_\beta(0,0))$

If $Z_{Bt} = 1$ then $V_\beta(0,1) = (h_0 + \Psi + \lambda_B\delta V_\beta(0,0))$.

Hence, $V_\beta(0,0) = \frac{x^* + \Psi + \lambda_B\delta p(h_0 + \Psi)}{(1-\lambda_B\delta)(p\lambda_B\delta + 1)}$.

If $1 < Z_{Bt} \leq \overline{Z}$ then $V_\beta(0, Z_{Bt}) = h_0 - c_B + \Psi + \delta\lambda_B(V_\beta(0, Z_{Bt}-1))$

$V_\beta(0,1) = h_0 + \Psi + \delta\lambda_\iota(V_\beta(0,0))$

$V_\beta(0,2) = h_0 - c_B + \Psi + \delta\lambda_B(V_\beta(0,1)) = h_0 - c_B + \Psi + \delta\lambda_B(h_0 + \Psi + \delta\lambda_B(V_\beta(0,0)))$

More generally, if $Z_{Bt} = Z \leq \overline{Z}$ then $V_\beta(0,Z) = \delta^Z\lambda^Z V_\beta(0,0) + h_0\delta^{Z-1}\lambda^{Z-1} + \sum_{\tau=0}^{Z-2}\lambda^\tau\delta^\tau(h_0 - c_B) + \Psi\frac{1-\delta^Z\lambda^Z}{1-\delta\lambda_B}$

If $Z_{Bt} > \overline{Z}$ then $V_\beta(0, Z_{Bt}) = 1 - c_B + \Psi + \delta\lambda_B V_\beta(0, Z_{Bt}+1)$ since $V_\beta(0, Z_{Bt}) = V_\beta(0, Z_{Bt}+1)$ then $V_\beta(0, Z_{Bt}) = \frac{1-c_B+\Psi}{1-\delta\lambda_B}$

Continuation Values for citizens' in nation A. $V_A(Z_{At} = 1, .) = 0 + (1-\lambda_A)\delta V_A(0,0)$

$V_A(0,0) = (1-x^*) + R + \delta((1-\lambda_B p)V_A(0,0) + p\lambda_B V_A(0,1))$

$V_A(0,1) = (1-h_0) + R + \delta V_A(0,0)$

Substituting into the expression for $V_A(0,0)$ yields

$V_A(0,0) = \frac{1-x^*+\lambda_B\delta p(1-h_0)+R(1+\delta\lambda_B p)}{(1-\delta)(\lambda_B\delta p+1)}$

$V_A(0, 1 < Z_{Bt} \leq \overline{Z}) = (1-h_0) - c_A + R + \delta(\lambda_B V_\alpha(0, Z_{Bt}-1) + (1-\lambda_B)V_A(0,0))$.

If $Z_{Bt} > \overline{Z}$ then $V_A(0, Z_{Bt}) = -c_A + R + \delta(\lambda_B V_A(0, Z_{Bt}+1) + (1-\lambda_B)V_A(0,0))$ where $V_\alpha(0, Z_{Bt}+1) = V_\alpha(0, Z_{Bt})$ so if $Z_{Bt} > \overline{Z}$ then $V_A(0, Z_{Bt}) = \frac{R-c_A+\delta(1-\lambda_B)V_A(0,0)}{1-\lambda_B\delta}$.

Continuation Values for citizens' in nation B. If $Z_{At} = 1$ then $V_B(1,.) = 1 + \delta(\lambda_A V_B(1,0) + (1-\lambda_A)V_B(0,0))$

so $V_B(1,.) = \frac{1+\delta((1-\lambda_A)V_B(0,0))}{(1-\delta\lambda_A)}$

If $Z_{At} = 0$ and $Z_{Bt} = 0$ then

$V_B(0,0) = x^* + \delta(p\lambda_B V_B(0,1) + (1-p\lambda_B)V_B(0,0))$

If $Z_{Bt} = 1$ then $V_\beta(0,1) = h_0 + \delta V_B(0,0)$

Substitution in $V_B(0,1)$ yields $V_B(0,0) = \frac{x^*+h_0\lambda_B\delta p}{(1-\delta)(\lambda_B\delta p+1)}$.

If $1 < Z_{Bt} \leq \overline{Z}$ then

$V_B(0, Z_{Bt}) = h_0 - c_B + \delta(\lambda_B V_B(0, Z_{Bt}-1) + (1-\lambda_B)V_B(0,0))$

$V_B(0,1) = h_0 + \delta V_B(0,0)$

$V_2 = h_0 - c_B + \delta(\lambda_B V_1 + (1-\lambda_B)V_0)$

$V_3 = h_0 - c_B + \delta(\lambda_B V_2 + (1-\lambda_B)V_0)$

In general for $Z_{Bt} = Z \leq \overline{Z}$,

$V_B(0,Z) = V_B(0,0)\delta(\lambda_B^{Z-1}\delta^{Z-1} + \sum_{j=1}^{Z-1}(1-\lambda_B)\lambda_B^{j-1}\delta^{j-1}) + h_0\sum_{j=1}^{Z}\delta^{j-1}\lambda_B^{j-1} - c_B\sum_{j=1}^{Z-1}\delta^{j-1}\lambda_B^{j-1}$

If $Z_{Bt} > \overline{Z}$ then $V_B(0, Z_{Bt}) = 1 - c_B + \delta(\lambda_B V_B(0, Z_{Bt}+1) + (1-\lambda_B)V_B(0,0))$ and since if $Z_{Bt} > \overline{Z}$ then $V_B(0, Z_{Bt} > \overline{Z}) = V_B(0, Z_{Bt}+1)$ therefore if $Z_{Bt} > \overline{Z}$ then $V_B(0, Z_{Bt}) = \frac{1-c_B+\delta(1-\lambda_B)V_B(0,0)}{1-\delta\lambda_B}$.

Having characterized the continuation values we show that given the conditions stated each players strategy is utility maximizing for every possible state.

Leader α's decision to sanction B. Suppose $Z_{At} = 1$. Sanctions have no influence on future behavior so A never sanctions. Suppose $Z_{At} = 0$.

Case1) $Z_{Bt} = 0$. If α sanctions having observed y_{Bt} then α's payoff is

$U_\alpha(\text{sanction}|Z_{At} = 0, Z_{Bt} = 0) = 1 - x - c_A + \Psi + R + \lambda\delta(p(x^*)V_\alpha(0,1) + (1 - p)V_\alpha(0,0))$. If A does not sanction then

$U_\alpha(\text{no sanction}|Z_{At} = 0, Z_{Bt} = 0) = 1-x+\Psi+R+\lambda\delta(p(x^*)V_\alpha(0,1)+(1-p)V_\alpha(0,0))$. Since $c_A > 0$, A does not sanction unless sanctions are called for. Similarly, if $Z_{At} = 0$, $Z_{Bt} = 1$ and $y_t \leq h_0$ then A does not sanction.

Case 2) If $Z_{Bt} = 1$ and $y_{Bt} > h_0$ then $U_\alpha(\text{sanction}|Z_{At} = 0, Z_{Bt} = 1, y_{Bt} > h_0) = 1 - y_{Bt} - c_A + \Psi + R + \lambda_A\delta V_\alpha(0,2)$. $U_\alpha(\text{no sanction}|Z_{At} = 0, Z_{Bt} = 1, y_{Bt} > h_0) = 1-y_{Bt}+\Psi+R+\lambda_A\delta V_\alpha(1,0)$. Hence α sanctions providing $\lambda_A\delta[V_\alpha(0,2)-V_\alpha(1,0)] \geq c_A$

Case 3) If $Z_{At} = 0, Z_{Bt} > 1$ then $U_\alpha(\text{sanction}|Z_{At} = 0, Z_{Bt}, y_{Bt} > h_0) = 1 - y_{Bt} - c_A + \Psi + R + \lambda_A\delta V_\alpha(0, Z_{Bt} + 1)$.

$U_\alpha(\text{sanction}|Z_{At} = 0, Z_{Bt}, y_{Bt} \leq h_0) = 1 - y_{Bt} - c_A + \Psi + R + \lambda_A\delta V_\alpha(0, Z_{Bt} - 1)$

$U_\alpha(\text{no sanction}|Z_{At} = 0, Z_{Bt} = 1, y_{Bt} > 0) = 1 - y_{Bt} + \Psi + R + \lambda_A\delta V_\alpha(1,0)$.

So α sanctions providing $\lambda_A\delta[V_\alpha(0, Z_{Bt} + 1) - V_\alpha(1,0)] \geq c_A$ if $y_{Bt} > h_0$ and $\lambda_A\delta[V_\alpha(0, Z_{Bt} - 1) - V_\alpha(1,0)] \geq c_A$ if $y_{Bt} \leq h_0$. The worst case scenario is that $Z_{Bt} > \overline{Z}$ in which case A sanctions with no prospect of restoring cooperation until leader β dies. Hence A sanctioning in all appropriate states implies that $c_A \leq \delta\lambda_A \frac{(\delta\lambda_A(1-\delta\lambda_A)(1-\lambda_B))V_\alpha(0,0)-\delta\lambda_A(1-\lambda_B)\Psi+(1-\delta\lambda_A)R}{(1-\delta\lambda_A)(1+\delta\lambda_A(1-\lambda_B))}$.

β's decisions given α's strategy. We show that leader β behaves optimally given the state and other players' strategies.

Case1: If $Z_{At} = 1$ then A never sanctions so $x_t = 1$.

Case 2: $Z_{At} = 0$.

2a) If $Z_{Bt} = 0$ then

$V_\beta(0,0) = x^* + \Psi + \lambda_B\delta(pV_\beta(0,1) + (1-p)V_\beta(0,0))$ and $V_\beta(0,1) = h_0 + \Psi + \lambda_B\delta(V_\beta(0,0))$ so $V_\beta(0,0) = \frac{x^*+\Psi+\lambda_B\delta p(h_0+\Psi)}{(1-\lambda_B\delta)(p\lambda_B\delta+1)}$. Maximizing this with respect to x* yields:

$\frac{d}{dx}V_\beta(0,0) = \frac{p(x)\lambda_B\delta+1+\lambda_B\delta\frac{\partial p(x)}{\partial x}h_0-\lambda_B\delta\frac{\partial p(x)}{\partial x}x}{(1-\lambda_B\delta)(p(x)\lambda_B\delta+1)^2} = 0$ so the FOC is $p(x)\lambda_B\delta + 1 + \lambda_B\delta\frac{\partial p(x)}{\partial x}(h_0-x) = 0$ where $\frac{dp}{dx} = \frac{d}{dx}(1-F(h-x)) = F'(h-x)$ and $\frac{d^2}{dx^2}(1-F(h-x)) = -F''(h-x)$. The corresponding SOC evaluated at this point is

$$\frac{d^2}{dx^2}\left(\frac{x+\Psi+\lambda_B\delta p(x)(h_0+\Psi)}{(1-\lambda_B\delta)(p(x)\lambda_B\delta+1)}\right) =$$
$$-\lambda_B\delta\frac{-\frac{\partial^2 p(x)}{\partial x^2}h_0 p(x)\lambda_B\delta-\frac{\partial^2 p(x)}{\partial x^2}h_0+2\frac{\partial p(x)}{\partial x}p(x)\lambda_B\delta+2\frac{\partial p(x)}{\partial x}+2\lambda_B\delta\frac{\partial p(x)}{\partial x}^2 h_0-2\frac{\partial p(x)}{\partial x}^2\lambda_B\delta x+\frac{\partial^2 p(x)}{\partial x^2}xp(x)\lambda_B\delta+\frac{\partial^2 p(x)}{\partial x^2}x}{(1-\lambda_B\delta)(p(x)\lambda_B\delta+1)^3}$$

where the numerator is $\lambda_B\delta F'(h-x)F''(h-x)(x-h_0)^2$. Thus a sufficient condition for a maximum is that $F''(h-x) < 0$ which is satisfied by the concavity of the exponential distribution. Hence $x_{Bt} = x^*$ is an optimal choice given $Z_{Bt} = 0$.

Case 2b) If $Z_{Bt} = 1$ then $U_\beta(x_t \leq h_0|Z_{At} = 0, Z_{Bt} = 1) = x_t + \Psi + \lambda_B\delta V_\beta(0,0)$. This payoff is maximized by $x_t = h_0$. If $x_t > h_0$ then $U_\beta(x_t > h_0|Z_{At} = 0, Z_{Bt} = 1) = x_t - c_B + \Psi + \lambda_B\delta V_\beta(0,2)$, which is maximized by $x_t = 1$. Thus, β 'apologizes' if $h_0 + \lambda_B\delta V_\beta(0,0) \geq 1 - c_B + \lambda_B\delta V_\beta(0,2)$ where

$V_\beta(0,2) = h_0 - c_B + \Psi + \delta\lambda_B(h_0 + \Psi + \delta\lambda_B(V_\beta(0,0)))$ if $2 \leq \overline{Z}$ and $V_\beta(0,2) = \frac{1-c_B+\Psi}{1-\delta\lambda_B}$ otherwise. If $2 \leq \overline{Z}$ then β apologizes tomorrow. Since it is always better to apologize immediately then the binding constraint is $h_0 + \lambda_B\delta V_\beta(0,0) \geq 1 - c_B + \lambda_B\delta(\frac{1-c_B+\Psi}{1-\delta\lambda_B})$, which implies $c_B \geq \frac{(-\lambda_B\delta p-1+\lambda_B\delta)h_0+\lambda_B\delta p+1-\delta\lambda_B x^*}{\lambda_B\delta p+1}$.

More generally β apologizes iff $Z_{Bt} \leq \overline{Z}$, where $\overline{Z}$ is defined such that $V_\beta(0, Z_{Bt} = Z) = \delta^Z\lambda^Z V_\beta(0,0) + h_0\delta^{Z-1}\lambda^{Z-1} + \sum_{\tau=0}^{Z-2}\lambda^\tau\delta^\tau(h_0 - c_B) + \Psi\frac{1-\delta^Z\lambda^Z}{1-\delta\lambda_B} \geq \frac{1-c_B+\Psi}{1-\delta\lambda_B}$ for all $Z_{Bt} \leq \overline{Z}$ and the opposite is true for all $Z_{Bt} > \overline{Z}$.

Citizens' decision to depose leader in nation A If the citizens depose leader α then they pay cost r_A. If $Z_{At} = 0$ and not condition J then $U_A(\text{replace}|Z_{At} =$

$0, Z_{Bt}) = U_A(\text{retain}|Z_{At} = 0, Z_{Bt}) - r_A$, so the citizens retain their leader.

If $Z_{At} = 1$ or condition J then $U_A(\text{replace}|Z_{At} = 1, Z_{Bt}) = -r_A + \delta V_A(0,0)$ and $U_A(\text{retain}|Z_{At} = 0, Z_{Bt}) = \delta\lambda_A V_A(1,0) + \delta(1-\lambda_A)V_A(0,0)$, where $V_A(Z_{At} = 1, .) = (1-\lambda_A)\delta V_A(0,0)$, and $V_A(0,0) = \frac{1-x^*+\lambda_B\delta p(1-h_0)+R(1+\delta\lambda_B p)}{(1-\delta)(\lambda_B\delta p+1)}$. So the citizens retain their leader if $r_A \geq \delta\lambda_A(\delta\lambda_A - \delta + 1) V_A(0,0)$.

Citizens' decision to depose leader in nation B. If the citizens depose leader β then they pay cost r_B. If $Z_{At} = 1$ or condition J and $L_{At} = 0$ then $U_B(\text{replace}|Z_{At}, Z_{Bt} = 0) = -r_B + \delta V_B(1,0)$ and $U_B(\text{retain}|Z_{At}, Z_{Bt} = 0) = \delta V_B(1,0)$ so the citizens retain their leader.

Suppose $Z_{At} = 0$ and not condition J.

At the time the citizens make their decisions they know what the state will be at the beginning of the next period if they do not remove their leader. For notational convenience we denote this state as Z_{Bt+1}.

Suppose without leadership deposition tomorrow's state will be $Z_{Bt+1} = 0$ (either $Z_{Bt} = 0$ and $y_{Bt} \leq h$ or $Z_{Bt} = 1$ and $y_{Bt} \leq h_0$). $U_B(\text{replace}|Z_{At} = 0, Z_{Bt+1} = 0) = -r_B + \delta V_B(0,0)$ and $U_B(\text{retain}|Z_{At} = 0, Z_{Bt+1} = 0) = \delta V_B(0,0)$. Hence B retains their leader.

Suppose without leadership deposition tomorrow's state will be $Z_{Bt+1} = 1$ (Either $(Z_B = 0$ and $y > h)$ or $(Z_{Bt} = 2$ and $s_t = 1$ and $y_t \leq h_0))$.

$U_B(\text{replace}|Z_{At} = 0, Z_{Bt+1} = 1) = -r_B + \delta V_B(0,0)$ and $U_B(\text{retain}|Z_{At}, Z_{Bt+1} = 1) = \delta\lambda_B V_B(0,1) + \delta(1-\lambda_B)V_B(0,0)$. Thus B retains leader if $r_B \geq \lambda_B\delta(V_B(0,0) - V_B(0,1))$. More generally, if without replacement tomorrow's state will be Z then $U_B(\text{replace}|Z_{At} = 0, Z_{Bt}) = -r_B + \delta V_B(0,0)$ and $U_B(\text{retain}|Z_{At}, Z_{Bt}) = \delta\lambda_B V_B(0,Z) + \delta(1-\lambda_B)V_B(0,0)$. Hence in general the citizens retain β if $r_B \geq \lambda_B\delta(V_B(0,0) - V_B(0, Z_{Bt}))$ for all Z_{Bt}.

Given that all players actions are optimal given the strategies of the other players and the evolution of the state variable, equilibrium 1 is a MPE. QED.

Equilibrium 2.—

Here we formally state equilibrium 2 and show it is a MPE. We follow the same procedure used in equilibrium 1: first we calculate the continuation values for each player; second we show no utility improving defections exist for any player in any state. Where the proof is identical to that of equilibrium 1 we refer to the earlier proof.

Proposition 2: If the cost of removing leader α is low $(r_A \leq \delta\lambda_A \frac{1-\delta}{1-\lambda_A\delta} V_A(0,0))$, the cost of removing leader β is high $(r_B \geq \lambda_B\delta(V_B(0,0) - V_B(0, Z_{Bt}))$ for all $Z_{Bt})$, $c_A \leq \Psi + \lambda_A\delta\lambda_B R + \delta\lambda_A(1-\lambda_B)V_\alpha(0,0)$ and $c_B \geq \frac{(-\lambda_B\delta p - 1 + \lambda_B\delta)h_0 + \lambda_B\delta p + 1 - \delta\lambda_B x^*}{\lambda_B\delta p + 1}$ then there exists a Markov Perfect Equilibrium where

1) $x_{B,t} = \begin{cases} 1 & if \quad Z_{At} = 1 \text{ or } (Z_{At} = 0,\ Z_{Bt} > \overline{Z}) \\ h_0 & if \quad Z_{At} = 0,\ 0 < Z_{Bt} \leq \overline{Z} \\ x^* & if \quad Z_{At} = 0,\ Z_{Bt} = 0 \end{cases}$

2) $s_t = \begin{cases} 1 \quad if \quad Z_{At} = 0, \text{ and (either } Z_{Bt} > 1 \text{ or } (Z_{Bt} = 1 \text{ and } y_{B,t} > h_0)) \\ 0 \qquad \text{otherwise} \end{cases}$

3) $\gamma_A = \begin{cases} 1 \quad if \quad Z_{At} = 1 \text{ or condition } J \\ 0 \qquad \text{otherwise} \end{cases}$

4) $\gamma_B = 0.$

and x^* solves $p\lambda_B\delta + 1 + \lambda_B\delta\frac{dp}{dx}(h_0 - x^*) = 0$. Definitions used in the statement of the equilibrium are defined below.

Proof:

Continuation values. If $Z_{At} = 1$ then A loses its reputation externally, nation B plays $x_{Bt} = 1$ and the citizens depose leader α. Therefore, $V_\alpha(Z_{At} = 1, .) = 0$. Note we assume that leader α does not receive the benefits of the international game once removed from office. Leader α's continuation values for other states are the same as characterized in equilibrium 1.

Leader β: Since the citizens in A depose their leader, if $Z_{At} = 1$ then A's lost reputation lasts only one period: $V_\beta(1,.) = 1 + \Psi + \delta\lambda_B V_\beta(0,0)$. In other states leader β's continuation values are as in equilibrium 1.

Citizens' in nation A: If leader α has a bad reputation then the citizens remove her, so $V_A(Z_{At} = 1,.) = 0 - r_A + \delta V_A(0,0)$. The continuation values for other states are as in equilibrium 1.

Citizens' in nation B: If $Z_{At} = 1$ then $V_B(1,.) = 1 + \delta V_B(0,0)$. The continuation values for other states are as in equilibrium 1.

Given that $V_\beta(0,0)$ is identical to that of equilibrium and play on the equilibrium path is identical the FOC determining x^* is the same as in equilibrium 1. Indeed beyond α's decision to sanction and A's decision to remove their leader the decision are identical to equilibrium 1, hence we examine only those two decisions here.

A's decision to sanction. Case 1) Suppose $Z_{At} = 1$, $Z_{Bt} = 0$ or ($Z_{Bt} = 1$ and $y_{Bt} \leq h_0$) then sanctions have no influence on future behavior so A never sanctions.

case 2) $Z_{At} = 0$ and $Z_{Bt} = 1$.

2a) If $y_{Bt} > h_0$ then $U_\alpha(\text{sanction}|Z_{At} = 0, Z_{Bt} = 1, y_{Bt} > h_0) = 1 - y_{Bt} - c_A + \Psi + R + \lambda_A\delta(\lambda_B V_\alpha(0,2) + (1-\lambda_B)V_\alpha(0,0))$. $U_\alpha(\text{no sanction}|Z_{At} = 0, Z_{Bt} = 1, y_{Bt} > h_0) = 1 - y_{Bt} + R$. Therefore α sanctions providing $\Psi + \lambda_A\delta(\lambda_B V_\alpha(0,2) + (1-\lambda_B)V_\alpha(0,0)) \geq c_A$.

More generally in state $Z_{At} = 0$, $Z_{Bt} > 1$,

$U_\alpha(\text{sanction}|Z_{At} = 0, Z_{Bt}, y_{Bt} > h_0) = 1 - y_{Bt} - c_A + \Psi + R + \lambda_A\delta(\lambda_B V_\alpha(0, Z_{Bt} + 1) + (1-\lambda_B)V_\alpha(0,0))$, $U_\alpha(\text{sanction}|Z_{At} = 0, Z_{Bt}, y_{Bt} \leq h_0) = 1 - y_{Bt} - c_A + \Psi + R + \lambda_A\delta(\lambda_B V_\alpha(0, Z_{Bt} - 1) + (1-\lambda_B)V_\alpha(0,0))$ and $U_\alpha(\text{no sanction}|Z_{At} = 0, Z_{Bt} > 1) = 1 - y_{Bt} + R$.

Therefore α sanctions providing $\Psi + \lambda_A\delta(\lambda_B V_\alpha(0, Z_{Bt}+1) + (1-\lambda_B)V_\alpha(0,0)) \geq c_A$ if $y_{Bt} > h_0$ and $\Psi + \lambda_A\delta(\lambda_B V_\alpha(0, Z_{Bt} - 1) + (1-\lambda_B)V_\alpha(0,0)) \geq c_A$ if $y_{Bt} \leq h_0$. The worst case scenario is that $Z_{Bt} > \overline{Z}$ in which case A sanctions with no prospect of restoring cooperation until leader β dies, hence for α to sanction in all appropriate states requires $c_A \leq \Psi + \lambda_A \delta \lambda_B R + \delta \lambda_A (1-\lambda_B) V_\alpha(0,0)$.

Citizens' decision to depose leader in nation A. As in equilibrium 1 citizens do not remove their leader 'unnecessarily' since it imposes cost r_A. If condition J or $Z_{At} = 1$ then $U_A(\text{replace}|Z_{At} = 1, Z_{Bt}) = -r_A + \delta V_A(0,0)$ and $U_A(\text{retain}|Z_{At} = 0, Z_{Bt}) = \delta \lambda_A V_A(1,0) + \delta(1-\lambda_A) V_A(0,0)$, where $V_A(Z_{At} = 1, .) = -r_A + \delta V_A(0,0)$.

Therefore, if $Z_{At} = 1$ or condition J, then the citizens depose their leader if $r_A \leq \delta \lambda_A \frac{1-\delta}{1-\lambda_A \delta} V_A(0,0)$. Otherwise α is retained.

Since no player can improve their payoff by a one period deviation in any given state, equilibrium 2 is a MPE. QED.

Proof of Proposition 3..—

We start by characterizing continuation values in all states.

Continuation values of leader α. If $Z_{At} = 1$ then $V_\alpha(Z_{A,t} = 1, .) = 0$. If $Z_{At} = 0$ then $V_\alpha(0,0) = (1-x^*) + R + \Psi + \delta \lambda_A ((1-\lambda_B p) V_\alpha(0,0) + p \lambda_B V_\alpha(0,1))$. $V_\alpha(0,1) = (1-h_0) + R + \Psi + \delta \lambda_A (V_\alpha(0,0))$. Therefore, $V_\alpha(0,0) = \frac{(1+\lambda_A \delta \lambda_B p)(R+\Psi) + (1-x^*) + \lambda_A \delta \lambda_B p (1-h_0)}{(1-\lambda_A \delta)(1+\lambda_A \delta \lambda_B p)}$. If $Z_{Bt} = 2$ then $V_\alpha(0,2) = 1 - h_0 - c_A + R + \Psi + \delta \lambda_A ((1-\lambda_B) V_\alpha(0,0) + \lambda_B V_\alpha(0,1))$. If $Z_{Bt} > 2$ then $V_\alpha(0, Z_{Bt}) = -c_A + R + \Psi + \delta \lambda_A V_\alpha(0,0)$.

Continuation values of leader β. If $Z_{A,t} = 1$ then $V_\beta(1, .) = 1 + \Psi + \delta \lambda_B V_\beta(0,0)$. If $Z_{A,t} = 0$ and $Z_{B,t} = 0$ then $V_\beta(0,0) = x + \Psi + \lambda_B \delta (p V_\beta(0,1) + (1-p) V_\beta(0,0))$. $V_\beta(0,1) = h_0 + \Psi + \lambda_B \delta V_\beta(0,0)$. Therefore, $V_\beta(0,0) = \frac{x + \Psi + \lambda_B \delta p (h_0 + \Psi)}{(1 - \lambda_B \delta p \lambda_B \delta - \lambda_B \delta (1-p))}$. $V_\beta(0, Z_{Bt} = 2) = h_0 - c_B + \Psi + \delta \lambda_B V_\beta(0,1)$. $V_\beta(0, Z_{Bt} > 2) = 1 - c_B$.

Continuation values for citizens' in nation A. $V_A(1,.) = -r_A + \delta V_A(0,0)$. $V_A(0,0) = (1-x^*) + R + \delta((1-\lambda_B p)V_A(0,0) + p\lambda_B V_A(0,1))$. $V_A(0,1) = 1 - h_0 + R + \delta V_A(0,0)$. Substituting this into the expression for $V_A(0,0)$ yields $V_A(0,0) = \frac{\delta\lambda_B p(1-h_0)+\delta\lambda_B pR+1-x+R}{(1-\delta)(1+\delta\lambda_B p)}$. $V_A(0,2) = 1 - h_0 - c_A + R + \lambda_B \delta V_A(0,1) + \delta(1-\lambda_B)V_A(0,0)$. If $Z_{Bt} > 2$ then $V_A(0, Z_{Bt}) = -c_A + R + \delta V_A(0,0)$.

Continuation values for citizens' in nation B. If $Z_{A,t} = 1$ then $V_B(1,.) = 1 + \delta V_B(0,0)$. If $Z_{At} = 0$ and $Z_{Bt} = 0$ then $V_B(0,0) = x^* + \delta(p\lambda_B V_B(0,1) + (1 - p\lambda_B)V_B(0,0))$. $V_B(0,1) = h_0 + \delta V_B(0,0)$. Substitution yields $V_B(0,0) = \frac{x^*+h_0\delta\lambda_B p}{(1-\delta)(1+\delta\lambda_B p)}$. $V_B(0,2) = h_0 - c_B + \delta\lambda_B V_B(0,1) + \delta(1-\lambda_B)V_B(0,0)$. If $Z_{Bt} > 1$ then $V_B(0, Z_{B,t}) = 1 - c_B - r_B + \delta V_B(0,0)$.

Given these continuation values we show that each player's strategy is utility maximizing given the state and other players' strategies.

α's decision to sanction B. Suppose $Z_{At} = 1$, $Z_{Bt} = 0$, or ($Z_{Bt} = 1$ and $y_{Bt} \leq h_0$). Under these conditions sanctions have no influence on future behavior so A never sanctions since $c_A > 0$.

Suppose $Z_{At} = 0$ and either ($Z_{Bt} = 1$ and $y_{Bt} > h_0$) or $Z_{Bt} > 2$ or ($Z_{Bt} = 2$ and $y_{Bt} > h_0$) then $U_\alpha(\text{sanction}|.) = 1 - y_{Bt} - c_A + \Psi + R + \lambda_A \delta V_\alpha(0,0)$. $U_\alpha(\text{no sanction}|.) = 1 - y_{Bt} + R$. Therefore α sanctions providing $\Psi + \lambda_A \delta V_\alpha(0,0) \geq c_A$.

If ($Z_{At} = 0$, $Z_{Bt} = 2$ and $y_{Bt} \leq h_0$) then $U_\alpha(\text{sanction}|.) = 1 - y_{Bt} - c_A + \Psi + R + \lambda_A \delta V_\alpha(0,1)$. $U_\alpha(\text{no sanction}|.) = 1 - y_{Bt} + R$. Therefore, α sanctions providing $\Psi + \lambda_A \delta V_\alpha(0,1) \geq c_A$. Hence α's decision to sanction is optimal.

β's decisions given α's strategy. Case1: $Z_{At} = 1$. Under this circumstance A never sanctions so β plays $x_t = 1$.

Suppose $Z_{At} = 0$, $Z_{Bt} = 0$. Leader β's maximization program generates the same FOC as in equilibrium 1.

Given citizen B's strategy, leader β has one opportunity to apologize. That is to say if the next period starts with $Z_{Bt+1} = 1$ then the voters will not remove the leader, but instead given her the opportunity to apologize. However, if the state variable starting in the next period means that two period of apology are need ($Z_{Bt+1} \geq 2$) then the citizens remove β immediately.

If $Z_{A,t} = 0$ and $Z_{B,t} = 1$ then $U_\beta(x_{Bt} \leq h_0|0,1) = x_{Bt} + \Psi + \delta\lambda_B V_\beta(0,0)$, which is maximized for $x_{Bt} = h_o$. $U_\beta(x_{Bt} > h_0|0,1) = x_{Bt} - c_B$, which is maximized for $x_{Bt} = 1$. Since $\Psi > 1$ then $c_B \geq 1 - h_0 - \Psi - \lambda_B\delta V_\beta(0,0)$ so β apologizes.

If $Z_{A,t} = 0$ and $Z_{B,t} = 2$ then $U_\beta(x_{Bt} \leq h_0|0,2) = x_{Bt} - c_B + \Psi + \delta\lambda_B V_\beta(0,1)$, which is maximized for $x_{Bt} = h_o$. $U_\beta(x_{Bt} > h_0|0,2) = x_{Bt} - c_B$, which is maximized for $x_{Bt} = 1$. Since $\Psi > 1$ then β apologizes.

If $Z_{A,t} = 0$ and $Z_{B,t} \geq 3$ then $U_\beta(x_{Bt}|0, Z_{Bt}) = x_{Bt} - c_B$, which is maximized for $x_{Bt} = 1$.

Hence β apologizes if apologizing will preserve her term in office. Otherwise she sets $x_{Bt} = 1$.

Citizens' decision to depose leader in nation A As in equilibrium 2 citizens do not remove their leader 'unnecessarily' since it imposes cost r_A. If condition J or $Z_{At} = 1$ then $U_A(\text{replace}|.) = -r_A + \delta V_A(0,0)$ and $U_A(\text{retain}|.) = \delta\lambda_A V_A(1,0) + \delta(1-\lambda_A)V_A(0,0)$, where $V_A(1,.) = -r_A + \delta V_A(0,0)$. Therefore, if $Z_{At} = 1$ or condition J, then the citizens depose their leader if $r_A \leq \delta\lambda_A \frac{1-\delta}{1-\lambda_A\delta} V_A(0,0)$. Otherwise α is retained.

Citizens' decision to depose leader in nation B..—

If the citizens depose leader β then they pay cost r_B. If the state in the next period will be $Z_{Bt+1} = 0$ (i.e. either $Z_{At} = 1$ or $Z_{Bt} = 0$ and $y_{Bt} \leq h$ or $Z_{Bt} = 1$ and $y_{Bt} \leq h_0$) then $U_B(\text{replace}|Z_{At}, Z_{Bt} = 0) = -r_B + \delta V_B(0,0)$ and $U_B(\text{retain}|Z_{At}, Z_{Bt} =$

$0) = \delta V_B(Z_{At}, 0)$ so the citizens retain their leader.

Suppose $Z_{At} = 0$. Suppose without leadership deposition tomorrow's state will be $Z_{Bt+1} = 1$ (i.e. $Z_{Bt} = 0$ and $y_{Bt} > h$ or $Z_{Bt} = 2$ and $y_{Bt} \leq h_0$). $U_B(\text{replace}|.) = -r_B + \delta V_B(0,0)$ and $U_B(\text{retain}|.) = \lambda_B \delta(h_0 + \delta V_B(0,0)) + \delta(1-\lambda_B)V_B(0,0)$. Hence B retains leader β if $r_B \geq \delta\lambda_B((1-\delta)V_B(0,0) - h_0)$.

Suppose without leadership deposition tomorrow's state will be $Z_{Bt+1} = 2$ (i.e. $Z_{Bt} = 3$ and $y_{Bt} \leq h_0$ or $Z_{Bt} = 1$ and $y_{Bt} > h_0$). $U_B(\text{replace}|.) = -r_B + \delta V_B(0,0)$ and $U_B(\text{retain}|.) = \lambda_B \delta(h_0 + \delta\lambda_B V_B(0,1) + \delta(1-\lambda_B)V_B(0,0)) + \delta(1-\lambda_B)V_B(0,0)$. Hence B deposes leader β if $r_B \leq -\delta\lambda_B\left(1+\delta\lambda_B\right)h_0 + \lambda_B\delta\left(1-\delta\right)\left(\delta\lambda_B+1\right)V_B(0,0)$.

Suppose without leadership deposition tomorrow's state will be $Z_{Bt+1} > 2$ (i.e. $Z_{Bt} > 3$ or ($Z_{Bt} = 2$ and $y_{Bt} > h_0$)). $U_B(\text{replace}|.) = -r_B + \delta V_B(0,0)$ and $U_B(\text{retain}|.) = \delta\lambda_B(1 - c_B - r_B + \delta V_B(0,0)) + \delta(1-\lambda_B)V_B(0,0)$. Thus B replaces leader if $r_B \leq \delta\lambda_B \frac{-1+c_B+(1-\delta)V_B(0,0)}{1-\delta\lambda_B}$.

Given that all players actions are optimal given the strategies of the other players and the evolution of the state variable, equilibrium 3 is a MPE. QED.

Proof of Proposition 4..—

We start by characterizing the continuation values associated with the strategy profile.

Continuation values of leader α. $V_\alpha(Z_{A,t} = 1, .) = 0$. $V_\alpha(0,0) = (1-x^*) + R + \Psi + \delta\lambda_A V_\alpha(0,0)$. Therefore, $V_\alpha(0,0) = \frac{\Psi+R+(1-x^*)}{1-\delta\lambda_A}$. $V_\alpha(0,1) = 1 - h_0 + R + \Psi + \delta\lambda_B V_\alpha(0,0)$. $V_\alpha(0, Z_{Bt} > 1) = -c_A + R + \Psi + \delta\lambda_A V_\alpha(0,0)$.

Continuation value for leader β. If $Z_{At} = 1$ then $V_\beta(1,.) = 1 + \Psi + \delta\lambda_B V_\beta(0,0)$

If $Z_{At} = 0$ and $Z_{Bt} = 0$ then $V_\beta(0,0) = x^* + (1-p)\Psi + (1-p)\lambda_B\delta V_\beta(0,0)$, so $V_\beta(0,0) = \frac{x^*+\Psi(1-p)}{1-\lambda_B\delta+\lambda_B\delta p}$. $V_\beta(0,1) = h_0 + \Psi + \delta\lambda_B V_\beta(0,0)$. $V_\beta(0, Z_{Bt} > 1) = 1 - c_A$.

Continuation values for citizens' in nation A. $V_A(1,.) = -r_A + \delta V_A(0,0)$. $V_A(0,0) = (1-x^*) + R + \delta V_A(0,0)$, so $V_A(0,0) = \frac{(1-x^*)+R}{1-\delta}$. $V_A(0,1) = 1 - h_0 + R + \delta V_A(0,0)$. If $Z_{Bt} > 1$ then $V_A(0, Z_{Bt}) = -c_A + R + \delta V_A(0,0)$.

Continuation values for citizens' in nation B. If $Z_{A,t} = 1$ then $V_B(1,.) = 1 + \delta V_B(0,0)$. If $Z_{A,t} = 0$ and $Z_{B,t} = 0$ then $V_B(0,0) = x^* - r_B p + \delta V_B(0,0)$, so $V_B(0,0) = \frac{x^* - r_B p}{1-\delta}$. $V_B(0,1) = h_0 + \delta V_B(0,0)$. If $Z_{Bt} > 1$ then $V_B(0, Z_{Bt}) = 1 - c_A - r_B + \delta V_B(0,0)$.

A's decision to sanction B. Suppose $Z_{At} = 1$ or ($Z_{Bt} = 0$ and $y_{Bt} \leq h_0$) then sanctions have no influence on future behavior so A never sanctions.

Suppose $Z_{At} = 0$ and either ($Z_{Bt} = 1$ and $y_{Bt} > h_0$) or $Z_{Bt} > 1$ then $U_\alpha(\text{sanction}|.) = 1 - y_{Bt} - c_A + \Psi + R + \lambda_A \delta V_\alpha(0,0)$. $U_\alpha(\text{no sanction}|.) = 1 - y_{Bt} + R$. Therefore α sanctions providing $c_A \leq \Psi + \lambda_A \delta V_\alpha(0,0)$, where $V_\alpha(0,0) = \frac{\Psi + R + (1-x^*)}{1-\delta\lambda_A}$.

Leader β's decisions. If $Z_{At} = 1$ then A never sanctions whatever β does so β plays $x_t = 1$. Suppose $Z_{At} = 0$ and $Z_{Bt} = 0$. Leader β's payoff is $\frac{x+\Psi(1-p(x))}{1-\lambda_B\delta+\lambda_B\delta p(x)}$ which when maximized with respect to x yields the FOC: $\frac{d}{dx}V_\beta(0,0) = \frac{d}{dx}\frac{x+\Psi(1-p(x))}{1-\lambda_B\delta+\lambda_B\delta p(x)} = 0$, which implies $(1-\lambda_B\delta+\lambda_B\delta p) - (\Psi+\lambda_B\delta x)\frac{\partial p(x)}{\partial x} = 0$. The SOC is $\frac{d^2}{dx^2}\frac{x+\Psi(1-p(x))}{1-\lambda_B\delta+\lambda_B\delta p(x)} = \frac{N}{(1-\lambda_B\delta+\lambda_B\delta p(x))^3}$, where $N = 2\lambda_B\delta\left(\Psi+\lambda_B\delta x\right)\frac{\partial p(x)}{\partial x}^2 - 2\lambda_B\delta\left(1-\lambda_B\delta+\lambda_B\delta p(x)\right)\frac{\partial p(x)}{\partial x} - (\Psi+\lambda_B\delta x)(1-\lambda_B\delta+\lambda_B\delta p(x))\frac{\partial^2 p(x)}{\partial x^2}$. Evaluated at the FOC then $N < 0$.

If $Z_{A,t} = 0$ and $Z_{B,t} = 1$ then $U_\beta(x_{Bt} \leq h_0|0,1) = x_{Bt} + \Psi + \delta\lambda_B V_\beta(0,0)$ which is maximized for $x_{Bt} = h_0$. $U_\beta(x_{Bt} > h_0|0,1) = x_{Bt} - c_B$, which is maximized for $x_{Bt} = 1$. Hence $x_{Bt} = h_0$ iff $1 - c_B \leq h_0 + \Psi + \delta\lambda_B V_\beta(0,0)$ which implies $1 - h_0 - \Psi - \lambda_B\delta V_\beta(0,0) \leq c_B$. Suppose $Z_{Bt} > 1$ then $U_\beta(x_{Bt}|0,1) = x_{Bt} - c_B$ which is maximized by $x_{Bt} = 1$.

Citizens' decision to depose leader α in nation A. As in equilibrium 1 citizens do not remove their leader 'unnecessarily' since it imposes cost r_A.

If condition J or $Z_{At} = 1$ then $U_A(\text{replace}|.) = -r_A + \delta V_A(0,0)$ and $U_A(\text{retain}|.) = \delta\lambda_A V_A(1,0) + \delta(1-\lambda_A)V_A(0,0)$, where $V_A(1,.) = -r_A + \delta V_A(0,0)$, and $V_A(0,0) = \frac{(1-x^*)+R}{1-\delta}$. Therefore, if $Z_{At} = 1$ or condition J, then the citizens depose their leader if $r_A \leq \delta\lambda_A \frac{1-\delta}{1-\lambda_A\delta} V_A(0,0)$. Otherwise α is retained.

Citizens' decision to depose leader in nation B..—

If the citizens depose leader β then they pay cost r_B. If $Z_{Bt} = 0$ and $y_{Bt} \leq h$ then $U_B(\text{replace}|Z_{At}, Z_{Bt} = 0) = -r_B + \delta V_B(Z_{At}, 0)$ and $U_B(\text{retain}|Z_{At}, Z_{Bt} = 0) = \delta V_B(Z_{At}, 0)$ so the citizens retain their leader. Similarly if $Z_{At} = 1$ then there is no need to replace leader.

Suppose $Z_{At} = 0$. Suppose without leadership deposition tomorrow's state will be $Z_{Bt+1} = 1$. $U_B(\text{replace}|.) = -r_B + \delta V_B(0,0)$ and $U_B(\text{retain}|.) = \lambda_B(\delta h_0 + \delta^2 V_B(0,0)) + (1-\lambda_B)\delta V_B(0,0)$. Thus B replaces leader if $y_{Bt} > h$, $Z_{Bt} = 0$ and $r_B \leq -\lambda_B\delta h_0 + \delta\lambda_B(1-\delta)V_B(0,0)$.

Suppose without leadership deposition tomorrow's state will be $Z_{Bt+1} > 1$. $U_B(\text{replace}|.) = -r_B + \delta V_B(0,0)$ and $U_B(\text{retain}|.) = \lambda_B(\delta(1-c_B-r_B) + \delta^2 V_B(0,0)) + (1-\lambda_B)\delta V_B(0,0)$. Thus B replaces leader if either ($y_{Bt} > h_0$ and $Z_{Bt} = 1$) or ($Z_{Bt} > 1$) and $r_B \leq \frac{\lambda_B\delta}{1-\lambda_B\delta}(-1 + c_B + (1-\delta)V_B(0,0))$.

Therefore, each player's decisions are optimal given the state and the other players strategy. Therefore, equilibrium 4 is a MPE. QED.

REFERENCES

Alt, James E., Randall L. Calvert, and Brian D. Humes. 1988. "Game Theory and Hegemonic Stability: The Role of Reputation and Uncertainty" *American Political Science Review*, 82(2): 445-66.

Axelrod, Robert and Robert O. Keohane. 1986. "Achieving Cooperation Under Anarchy: Strategies and Institutions." In *Cooperation Under Anarchy. ed.* Kenneth A. Oye. Princeton NJ: Princeton University Press.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson and James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge, MA: MIT Press.

Chayes, Abram and Antonia Handler Chayes. 1995. T*he New Sovereignty: Compliance with International Regulatory Agreements.* Cambridge MA: Harvard University Press.

Chiang, Alpha C. 2000. *Elements of Dynamic Optimization*. Prospect Heights IL: Waveland Press.

Downs, George W. and David M. Rocke. 1995. *Optimal Imperfection?: Domestic Uncertainty and Institutions in International Relations.* Princeton NJ: Princeton University Press.

Downs, George W. David M. Rocke and Peter Barsoom. 1996. "Is the Good News about Compliance Good News for Cooperation?" *International Organization.* 50(2): 379-406.

Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88(3): 577-592.

Fearon, James D. and David D. Laitin. 1996. "Explaining Interethnic Cooperation." *American Political Science Review* 90(4): 715-735.

Fundenberg, Drew and Jean Tirole. 1991. Game Theory. Cambridge MA: The MIT Press.

Gilligan, Michael J. 2004. "Is there a broader-deeper tradeoff in multilateral international agreements?" *International Organization* 58(3): 459-484.

Gilpin, Robert. 1981. *War and Change in World Politics*, New York: Cambridge University Press.

Goemans, Hein. E. 2000. *War and Punishment.* Princeton NJ: Princeton University Press.

Green, E. and R. Porter. 1984. "Non-cooperative Collusion under Imperfect Price Information." *Econometrica* 52(1): 87-100.

Guisinger, Alexandra and Alastair Smith. 2002. "Honest Threats: The Interaction of Reputation and Political Institutions in International Crises." *Journal of Conflict Resolution* 46(2): 175-200.

Hathaway, Oona. 2002. "Do Human Rights Treaties Make a Difference?" *Yale Law Journal* 111(8): 1935-2042.

Hathaway, Oona. 2004. Between Power and Principle: A Political Theory of International Law. Mimeo, Yale University Law School.

Hinojosa, Victor J. 2002. "Two-Level Games and International Narcotics Control: Cooperation Between the United States and Mexico and Columbia, 1997-2000." Paper presented at the annual meeting of the American Political Science Association, Boston MA.

Keohane, Robert.O. 1984. *After Hegemony.* Princeton, NJ: Princeton University Press.

Kindleberger, Charles. 1986. *The World In Depression 1929-1939.* Berkeley CA: University of California Press.

Kreps, David M. and Robert Wilson. 1982. Reputation and Imperfect Information. *Journal of Economic Theory* 27(2): 253-79.

Leeds, Brett Ashley. 1999. "Domestic Political Institutions, Credible Commitments and International Cooperation." *American Journal of Political Science* 43(4): 979-1002.

McCubbins, Mathew and Thomas Schwartz. 1987. "Congressional Oversight Overlooked: Police Patrols Versus Fire Alarms." In *Congress: Structure and Policy* eds., Mathew McCubbins and Terry Sullivan. New York: Cambridge University Press, pp. 426-40..

McGillivray, Fiona., and Alastair Smith. 2000. "Trust and Cooperation Through Agent-Specific Punishments" *International Organization* 54(4): 584-607.

McGillivray, Fiona., and Alastair Smith. 2004a. "The Impact of Leadership Turnover and Domestic Institutions on International Cooperation in a Noisy World." Working Paper, New York University.

McGillivray, Fiona., and Alastair Smith. 2004b. "The Impact of Leadership Turnover on Trading Relations Between States." *International Organizations* 58(3): 567-600.

McGillivray, Fiona., and Allan Stam. 2004. "Political Institutions, Coercive Diplomacy and the Duration of Sanctions." *Journal of Conflict Resolution.* 48(2):154-173. 2004

Porter R. 1983. "Optimal Cartel Trigger-price Strategies." *Journal of Economic Theory* 29(2): 313-38.

Raustiala, Kal and Slaughter, Anne-Marie. 2002. "International Law, International Relations and Compliance," The Handbook of International Relations Sage Publications, pp.538-58.

Rosendorff, B. Peter, and Helen V. Milner. 2001. "The Optimal Design of International Trade Institutions: Uncertainty and Escape." *International Organization* 55(4): 829– 857

Rosendorff, B. Peter. 2004. "Stability and Rigidity: Politics and Design of the WTO's Dispute Settlement Procedure." Working paper, May 2004.

Rieff, David. 2003. "Were Sanctions Right?" New York Times July 27, 2003 section 6, p 41.

Schultz, K. 2002. *Democracy and Coercive Diplomacy.* Cambridge: Cambridge University Press.

Selten, Richard. 1978. "The Chain-Store Paradox." *Theory and Decision* 9(2): 127-159.

Smith, Alastair. 1998. “International Crises and Domestic Politics.” *American Political Science Review.* 92(3): 623-638.