Rev. Econ. Design (2009) 13:251–278
DOI 10.1007/s10058-008-0054-7

ORIGINAL PAPER

Sanctions as revelation regimes

Daniel Verdier

Published online: 15 November 2008

Abstract Sanctioning in the face of uncertainty has been investigated from the perspective of signaling models, in which an informed target sends a signal and the sanctioner decides to sanction or not. As the more powerful power, however, the sanctioner could take the lead rather than react to the action of the target. Borrowing from contract theory, I present a new approach to the sanctioning of a target of unknown type in which the sanctioner explicitly structures the target's options before the target moves. The theory shows that in a situation in which a sanctioning state is ignorant of the target's willingness to comply, it may be able to elicit the revelation of that information through a careful mix of incentives. I derive the comparative statics and provide historical illustrations.

Keywords Sanction · International regime · Screening mechanism · Reward

JEL Classification F51

0 Introduction

Since the end of the Cold War, sanction regimes have become a ubiquitous tool of foreign policy in the hands of great powers seeking to maintain peace and order in

I thank Eric Fisher, Massimo Morelli, Chad Rector, Kuba Zielinski, and Byungwon Woo for valuable suggestions on prior versions.

D. Verdier (✉)
Department of Political Science, The Ohio State University,
2140 Derby Hall, 154 N. Oval Mall, Columbus, OH 43210-1373, USA
e-mail: verdier.2@osu.edu

the less developed parts of the world.[1] Still, the sanction literature laments the fact that imposed sanctions rarely elicit compliance (Hufbauer et al. 1990; Doxey 1996; Morgan and Schwebach 1997; Dashti-Gibson et al. 1997; Pape 1997; Drury 1998). A lot of the dissatisfaction with sanction regimes comes from the assumption that a sanction that elicits no voluntary compliance is a failed sanction. However, not all cases of defiance are cases of failure. In some circumstances, non-compliance is part of the sanctioner's optimal strategy. These circumstances primarily occur when the sanctioner is uncertain about the target's cost of compliance.

The claim that non-compliance is part of the sanctioner's optimal strategy in the face of uncertainty is not new. It has been investigated from the perspective of signaling models, in which an informed target sends a signal and the uninformed sanctioner decides to sanction or not. However, as the more powerful power, one might think that the sanctioner would take the lead rather than wait for the target to move. This article presents a new model of sanctioning with a target of unknown type in which a sanctioner explicitly structures the target's options before the target moves. I show that for some parametric configurations, the sanctioner is able to force information revelation through a carefully designed mix of sanctions and positive incentives. I call this strategy a "revelatory sanction regime." It differs from a "blind compliance regime," in which the sanctioner seeks to elicit compliance irrespective of target's resolve.

The model shows that revelation is an efficient strategy in situations of high uncertainty, while blind compliance is preferred when the sanctioner exhibits a strong preference for compliance. The model also points to a residual case in which the sanctioner's goals fall short of compliance and lead her to make an offer that the target is sure to reject. I call this strategy blind defiance or "sham offer." Sham offers may obtain when the sanctioner is less resolute than the least resolute target type.

I illustrate the empirical validity of these claims with three exemplary cases of sanction: revelation with the U.S. policy of engagement toward Iraq (1989); blind compliance with the U.S.'s threat to sanction a nuclear South Korea (1976), and a sham offer with the League's sanction of Italy (1935). I conclude by detailing the relevance of the model to extant policy debates on the efficiency of international sanctions. First, I detail the literature on sanctioning that is relevant to my argument.

1 The literature

This survey is limited to approaches that treat states as rational unitary actors, for it is for such rationalist explanations that the often-noted failure of imposed sanctions to elicit compliance is most puzzling.[2] The issue of whether sanctions do or should cause compliance has received three responses in the literature. First, some believe

[1] A study by the National Association of Manufacturers (1997) found that 35 countries were targeted by U.S. unilateral sanctions in the 1993-96 period.

[2] Imposing sanctions that fail to elicit compliance is less of a challenge for arguments featuring leaders or interest groups that can enjoy various side-effects of sanctioning without suffering costs (see Kaempfer and Lowenberg 1988).

that sanctions should be followed by compliance all the time but are not.[3] This more traditional literature devotes its attention to determining under what circumstances sanctions may work.[4]

Second, others, to whom I refer as the resource denial literature, argue that sanctions should and do cause compliance at least in part if not in full.[5] They claim that sanctions force compliance on a defiant target by working to subvert its government or denying it the very resources that it needs to sustain its harmful behavior. Sanctions, according to these authors, may or may not work at the threat stage by deterring defiance, but they definitely work, albeit partially, at the sanction stage by forcing compliance.

Third, still others argue that imposed sanctions are neither intended nor likely to cause compliance (Smith 1996; Miers and Morgan 1999; Schwebach 2000; Drezner 2003; Lacy and Niou 2004). These authors assume that sanctioning occurs under conditions of incomplete information about the target's cost of compliance or the sanctioner's degree of resolve or both, for, if both sides had complete information, they would always prefer to reach a negotiated solution rather than impose sanctions that are costly to all (Wagner 1988, p. 478). Sanctioning involves two stages: threat and implementation. Sanctions work, if at all, at the threat stage; sanctions that are actually imposed are sanctions that failed to work at the threat stage and are unlikely to work at the implementation stage.

To reduce uncertainty, the third literature privileges ex ante information revelation through signaling. In a signaling model the informed party moves before the uninformed party, with the former potentially giving out cues to the latter (Miers and Morgan 1999; Schwebach 2000; Lacy and Niou 2004). In the model by Lacy and Niou (2004), "compliant" and "resilient" target types separate in response to sender's sanctions in most equilibria. This is a fundamental result in which the sanctioner learns the type of the target by observing it play. The idea of using a regime to signal one's type is also developed by Kydd (2001), though his application is to joining international institutions rather than imposing a sanction regime.

I pick up the baton from where signaling models have taken it so far. Rather than having the informed target move first, I design a screening game in which the uninformed sanctioner sets the agenda for the informed target before this one moves. Having the sanctioner move first underscores the fact that sanctioning preeminently is about the deliberate and asymmetric designing of an institution—a sanction regime. Most sanction cases indeed do feature a big power seeking to alter the policy of a smaller country by means of a basket of positive or negative incentives.

[3] Hufbauer et al. (1990) studied 116 incidents involving economic sanctions and found a success rate of 34 percent.

[4] Dashti-Gibson et al. (1997) and Drury (1998) argue that inflicting insufficient pain on the target is a cause of failure. Also, sanctions are found not to work when imposed on agents that (1) have no significant trade links with the principal (Dashti-Gibson et al. 1997, p. 609; Davidson and Shambaugh 2000, p. 41; Nooruddin 2002 found otherwise), (2) are too powerful (Morgan and Schwebach 1997; Drury 1998), (3) are non-democratic (Nooruddin 2002), or (4) are long-term rivals who worry about the long-term consequences of acquiescing (Davidson and Shambaugh 2000; Drezner 1999; Nooruddin 2002 found otherwise).

[5] See Baldwin (1985) and Crawford (1999). See also Nooruddin (2002) for systematic evidence and Lopez and Cortright (2004) for qualitative evidence.

To explore the design component of sanction regimes, I borrow my tools from contract theory.[6] The sanction game presented here builds and expands on this benchmark in two ways. First, while contract theory does not deal with sanctions, the sanction game mixes carrots and sticks. Second, while contract theorists privilege the revelation of private information, I give the sanctioner the additional option of bypassing revelation altogether and attempting either to elicit universal compliance across target types or to force universal defiance. Usually dismissed as inefficient in contract theory, these two additional options can be efficient in sanction games.

I now present the model, starting with assumptions.

2 Assumptions

The game features a sanctioner who seeks a prize from a target. The sanctioner offers a package of reward and sanction without knowing the target's cost of compliance, which can be low or high. Upon receiving the offer, the target can reject any form of deal and declare war upon the sanctioner, in which case war ensues and the game ends. Or the target can accept to negotiate with the sanctioner and then choose between accepting or turning down the offer on the table. In either case, the game ends and promises and threats implemented.

I assume that the prize is indivisible at the bargaining stage, but divisible at the punishment stage. The assumption of indivisibility better captures what is a stake in sanctioning than would the reverse assumption of continuity. For instance, the West requested that Saddam ends the genocide of the Kurds, not that he merely reduces it by a certain percentage. However, in keeping with the resource denial literature, I contemplate the possibility that compliance may be partial at the punishment stage. The no-fly zone reduced the number of Kurds that were massacred by Saddam's military. The actual implementation of the sanction threat denied him critical resources, thereby yielding forced compliance, albeit partial. I also assume that the return on forced compliance should not be so high that it always pays off to sanction a target.

I define a sanction as any action that has a punitive effect short of war. This is an arbitrary limit, for in some circumstances, war is used as a sanction, while in others, sanctions are used in lieu of war. In this article, I restrict my attention to the case where the sanctioner tries to use sanctions in lieu of war, reserving the alternate case, where war is a sanction, to another endeavor.[7] The war aversion assumption may reflect exorbitant military costs or strategic considerations—for instance, Britain and France wanted to give the League of Nations the opportunity to sanction Mussolini in relation to the 1935 Italian invasion of Ethiopia but not start a war, let alone topple the fascist regime. However, despite the assumption that the sanctioner shirks war, I allow war

[6] The typical monopolistic screening game of contract theory features a transaction between a buyer and a seller where the seller does not know exactly how much the buyer is willing to pay for a good. The seller skirts the difficulty by offering the buyer a set of choices (each one saying "buy quantity z in exchange for sum of money t (z)") from which the buyer picks the option that maximizes her utility, thereby revealing her type and dissipating the initial uncertainty (see Laffont and Martimort 2002).

[7] See Verdier (2008). I do not consider the case in which sanctions are used to demonstrate the sanctioner's will to resort to war either. For a signaling game along such lines (see Schwebach 2000).

to be an option for the target, which may react to a sanction by declaring war on the sanctioner. For instance, Japan attacked Pearl Harbor because it preferred war to U.S. trade sanctions (Russett 1969). The war option provides the target with a reservation value and the sanctioner with an upper bound to sanction threats.

A potential problem is the presumption that the player who offers the package of positive and negative incentives can commit to it ex ante. This assumption makes sense in domestic settings, where sanction regimes take the form of contracts that are enforceable by the courts, but not necessarily in international relations, where no similar enforcement power exists a priori. However, when it comes to sanctioning, especially the sanctioning by the United States of a smaller country like Iraq, the sanctioner has to deliver on threats and promises to maintain its reputation as sanctioner. I assume that threats and promises are self-enforcing by means of an unmodelled repeated play between the sanctioner and successive targets taking them one at a time, in which sanctioner has a sufficiently long time horizon to invest in reputation building.[8]

The assumptions are formalized as follows. Define Γ^B as the following Bayesian game between a sanctioner and a target. The two dispute the allocation of a prize, denoted by z, with $z \in [0, Z]$ and $Z \in \Re^+$ denoting full compliance. In exchange for z, the sanctioner can promise transfer t and/or threaten sanction s at marginal cost σ, with $t, s, \sigma \geq 0$. The sanctioner's utility for target's compliance is $u(z)$, which, for the sake of simplicity, I specify as the linear function $u(z) = az$, with a a positive constant. The sanctioner maximizes function

$$U(z, s) = az - \sigma s \tag{1}$$

in the case where she imposes a sanction, while function

$$U(z, t) = az - t \tag{2}$$

in the case where she gives a reward. (Sanctioner is a "she," target a "he"). Conversely, the target's utility is defined as

$$V(z, s) = -s - \kappa z \tag{3}$$

if sanctioned and as

$$V(z, t) = t - \kappa z \tag{4}$$

if rewarded, in both cases with κ belonging to the type set $\Theta_T = \{k, K\}$ and $0 < k < K$. The target type is drawn from a distribution with h percent "soft" (k) types and $1 - h$ percent "tough" (K) types. "Soft" and "tough" are mere labels for low- and high-compliance cost players. Target knows his type while sanctioner only knows probability distribution h.

[8] See the chain store game in Kreps and Wilson (1981).

At the bargaining stage, the sanctioner is asking for $z = Z$ and not for some fraction thereof. The actual implementation of the sanction threat yields forced compliance $z \leq Z$, with z a variable that is endogenously determined by a resource denial technology of the type $z(\kappa) = \begin{cases} \frac{\zeta}{\kappa} s \text{ for values of } s \leq \frac{\kappa Z}{\zeta} \\ Z \text{ for } s > \frac{\kappa Z}{\zeta} \end{cases}$, with $0 \leq \zeta < 1$ and $\kappa \in \{k, K\}$. The first line of the z function means that the sanctioner extracts compliance z from an unwilling type with marginal cost κ by imposing direct sanction s. How partial the compliance is depends on an index of efficiency of the resource denial technology, as captured by the value of parameter ζ. If $\zeta > 0$, the sanction works both as an ex ante threat and, in the case where the threat fails, ex post denial. Parameter ζ is capped below unity to keep compliance partial.[9] If $\zeta = 0$, the sanction only works as an ex ante threat, with no possible follow-up in case of failure. The z function is bounded upward at $z = Z$, with the result that for any s that is equal or higher than $\frac{\kappa Z}{\zeta}$, the sanction induces compliance level Z.

The expected return on forced compliance should not be so high that it always pays off to sanction a target. For this condition to be satisfied, the expected marginal cost of sanctioning the target must be greater than the expected marginal return on sanctioning; formally, $\sigma > a\left(h\frac{\zeta}{k} + (1-h)\frac{\zeta}{K}\right)$, with σ the marginal cost of sanctioning, a the marginal gain of compliance, $\frac{\zeta}{k}$ the marginal efficiency of sanctioning the soft type, $\frac{\zeta}{K}$ the marginal efficiency of sanctioning the tough type, and h the frequency of soft types.

The target's war payoff, denoted $-w_T$, with $w_T > 0$, places a realistic cap on how severe threatened sanctions can be. The sanctioner cannot threaten to execute an expected sanction payoff that is lower than the target's payoff for war, lest the target prefer war to being sanctioned.

The last assumption is that the sanctioner's payoff for war, $-w_S$, with $w_S > 0$, is negative. This is the condition that implements the assumption that sanctioner prefers sanctioning to war. Because I assume that target can declare war rather than accept the deal, it would be possible without this condition for sanctioner to engineer war by confronting target with alternatives worse than war. One can guard against this possibility by making the sanctioner's war payoff negative.

An action for sanctioner is a sanction regime, that is, a message that states "concede Z in exchange for transfer t or face sanction s and be forced to yield z." The sanctioner chooses the values of t and s (recall that Z is a constant and z a function of s), the set of which is

$$M = \left\{(t, s) \in R_+^2\right\}.$$

The game starts with Nature drawing the target type and revealing the result of the draw to the target; sanctioner only knows the probability distribution. Sanctioner then

[9] To see this, note that the sanctioner can extract full compliance either by offering the reward-only deal $t = \kappa Z$, or by offering t and threatening sanction s in case of non-compliance, in which case we have $t = \kappa Z - s$, or by only threatening a large sanction $s = \kappa Z$, which can also be rewritten as $Z = \frac{s}{\kappa}$. Multiplying the right-hand side with $\zeta \in (0, 1)$ makes compliance partial.

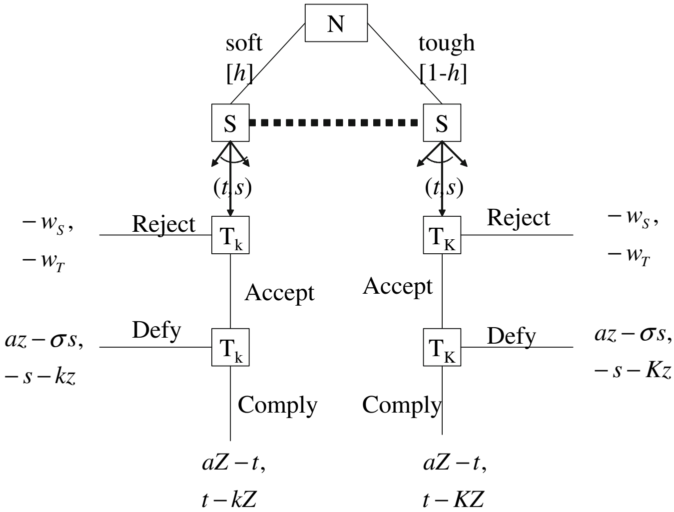

Fig. 1 The sanction regime game tree

proposes (t, s), which target can either reject or accept. If target rejects the offer, war occurs and war payoffs are allocated. If target accepts the offer, then target chooses to defy or to comply, with defiance triggering sanction payoffs and compliance inviting reward payoffs. The game tree is drawn in Fig. 1.

A strategy for sanctioner specifies the (t, s) menu she proposes. A strategy for target is the mapping $\Theta_T \times M \rightarrow \{R, A\}$ and $\Theta_T \times M \times \{A\} \rightarrow \{D, C\}$ specifying for each type and in response to all possible sanctioner' proposals two sequential responses: either accept the proposal or reject it and go to war; and, following accept, either comply or defy. The appropriate solution concept for the game is the Bayesian equilibrium, an equilibrium for which beliefs are exogenously given. An example of equilibrium is $\{(t, s)\,;\,(A, R; C, D)\}$, where sanctioner chooses a basket of positive and negative incentives, the soft target accepts and complies, and the tough target rejects and, off the equilibrium path, defies.

3 Definitions

Two definitions will help characterize and sort out the solutions of the game. The universe of offers, M, can be partitioned along two dimensions. A first dimension is the *compliance structure*, describing who complies, who defies, and who goes to war. A second is the *incentive structure*, capturing whether a reward, a sanction, or both are being proposed.

From the perspective of the compliance structure, the universe of proposals, M, can be partitioned into several subsets. In a first subset, the sanctioner chooses proposal (t, s) so that the target types choose to accept the proposal and pool on a common best reply. I call the kinds of message that enable this type of response "blind regimes," to reflect the fact that the sanctioner is still in the dark about the target's type when the game ends. The category blind regimes in turn subdivides into two subsets depending on whether the common reply is compliance or defiance. This allows me to define the following two subsets, a first, “blind compliance,” denoted $M^{BC} \in M$, in which the sanctioner chooses proposal (t, s) so that both types of target choose to accept the proposal and pool on compliance, and a second subset, “blind defiance,” denoted $M^{BD} \in M$, in which the sanctioner chooses proposal (t, s) so that the target types choose to accept the proposal and pool on defiance. I also refer to blind defiance proposals as “sham offers.”

In a third subset, denoted $M^{R} \in M$, the sanctioner selects proposal (t, s) so that the target types first accept the offer but then diverge in their response, with one type complying while the other defies. I call these messages “revelation regimes” to reflect the fact that information on the target’s type is eventually revealed.

I will refer to elements in M^{BC}, M^{BD}, and M^{R} alternately as “sanction regimes.” The residual set of proposals, $M^{W} \in M$, such that $M^{BC} \cup M^{BD} \cup M^{R} \cup M^{W} = M$, includes all proposals (t, s) so that at least one of the target types chooses war. This partitioning is formalized in Definition 1.

Definition 1 Define $M = M^{BC} \cup M^{BD} \cup M^{R} \cup M^{W}$ as the union of three types of sanction regimes, respectively blind compliance, blind defiance, and revelation, with the residual war scenario, such that

$$
\begin{aligned}
M^{BC} &= \{(t, s) \,|AC \in \arg\max V_{\kappa}(t, s), \text{ with } \kappa \in \{k, K\}\}, \\
M^{BD} &= \{(t, s) \,|AD \in \arg\max V_{\kappa}(t, s), \text{ with } \kappa \in \{k, K\}\}, \\
M^{R} &= \left\{(t, s) \left| \left\{ \begin{array}{l} AC \in \arg\max V_{k}(t, s) \\ AD \in \arg\max V_{K}(t, s) \end{array} \right\} \text{ or } \left\{ \begin{array}{l} AD \in \arg\max V_{k}(t, s) \\ AC \in \arg\max V_{K}(t, s) \end{array} \right\} \right.\right\},
\end{aligned}
$$

and M^{W} as including all the offers in which at least one of the target types prefers R to A.

Whether blind or revelatory, sanction regimes also differ with respect to their incentive structure. Given that there are two incentives, a positive and a negative, three kinds of regime are a priori conceivable: a sanction-only regime, in which the sanction threat alone yields whatever reply the sanctioner is seeking to elicit; a reward-only regime, in which it is the transfer that does the lifting; and a mixed regime, using both the threat and the promise. Hence the next definition.

Definition 2 A sanction-only regime is one in which $t = 0$ and $s > 0$, a reward-only regime, one in which $t > 0$ and $s = 0$, and a mixed regime, one in which $t > 0$ and $s > 0$.

Although the game can be solved at once, I proceed by type of regime for the sake of tractability and also to detail the institutional building component of the game. I solve the game first for the optimal menu $m \in M^{BC}$, then for the optimal $m \in M^{BD}$, and then for the optimal $m \in M^{R}$. I show that these solutions dominate all $m \in M^{W}$. I then have the sanctioner choose between her best response in M^{BC}, her best response in M^{BD}, and her best response in M^{R}.

4 Blind compliance regimes

Before describing the optimal blind compliance regime, it may be useful to provide a flavor of how the game is solved in Appendix. In a blind compliance regime, the sanctioner offers a proposal that does two things: (1) it incites both types of target to choose accept-and-comply and (2) it maximizes her utility. Deriving the solution thus takes two steps. In a first step, we establish the conditions that do just that—give an incentive to each target type to prefer accept-and-comply to any other move combination. Looking at the game tree, it is easy to see that for each type of target, two conditions have to be met. On the one hand, each type of target must prefer the accept-and-comply payoff to the reject payoff. On the other hand, each type of target must prefer the accept-and-comply payoff to the accept-and-defy payoff. These conditions, together with a few additional preconditions on the range of values that t and s are permitted to take, yield a list of constraints that the candidate equilibrium proposal must satisfy for blind compliance to obtain.

Once the universe of plausible candidate solutions has been thus narrowed down to those that meet all above-mentioned conditions, the next step involves finding the member of this selected universe that best satisfies these conditions, with "best" understood from the perspective of the sanctioner's utility. This utility is equal to $aZ - t$. This best proposal is calculated by means of a constrained maximization. Leaving further details to the Appendix, I state the solution to the optimization in Lemma 1, of which I then offer an intuitive interpretation.

Lemma 1 (Blind compliance) *The blind compliance regime that maximizes the sanctioner's utility is*:

- *the mixed menu* $\left(KZ - w_T, \frac{w_T}{1+\zeta} \leq s^*\right)$ *if* $0 < w_T < KZ$,
- *the sanction-only menu* $\left(0, \frac{KZ}{1+\zeta} \leq s^*\right)$ *if* $w_T \geq KZ$.

The first thing to notice is that both families of equilibria include a sanction threat, either in combination with a transfer (mixed menu) or alone (sanction-only menu). The reason for this is that there is always a mixed menu that beats the reward-only menu because sanctions are never imposed on the equilibrium path—both types comply—and are thus costless whereas rewards have to be delivered.

In a similar vein, a sanction-only menu will always beat a mixed menu, provided, of course, that a sanction-only menu be feasible. This will not always be the case, however. Recall the assumption that the sanctioner wants to avoid war and can only do so by offering the target a deal that is better than his war payoff, $-w_T$. This war payoff ranges between a number very close, yet inferior, to zero and negative infinity. The value that w_T can take, within that range, defines two cases. In a first case, w_T is very high (that is, the target's war payoff $-w_T$ is very low), thereby providing the sanctioner ample margin for eliciting compliance without having to offer a reward; this is the sanction-only menu. In this menu, the threatened sanction s can be very steep.

In a second case, w_T is not high enough to allow the sanctioner to rely solely on a sanction threat to elicit full compliance. The sanction threat must be supplemented with the promise of a reward. This is the mixed menu case, in which the sanctioner offers a reward that is equal to the tough target's cost of compliance KZ minus the sanction threat s.

The results reflect the particularity of blind compliance regimes: sanctions are never imposed on the equilibrium path while rewards have to be delivered. The sanctioner always prefers sanctions to rewards, explaining the absence of a reward-only menu. The only constraint on sanctioning comes from the target's war payoff, $-w_T$, as just explained. Therefore, blind compliance regimes can be very cheap if sanctions are feasible, but prohibitively expensive if sanctions have to be supplemented with a reward, for the reward has to be high enough to convince the tough type to comply.

A caveat is in order with respect to Lemma 1. The value of the sanction is indeterminate within a given range. Indeterminacy reflects the fact that sanctioning in an equilibrium where all target types comply is located off the equilibrium path. Sanctions are never imposed. They have to meet a minimum to be efficient, but the maximum is indeterminate. Admittedly unrealistic, this feature is of no concern to us for two reasons: it has no bearing on the results and it could be easily eliminated by attaching a cost, however small, to uttering a sanction threat.

5 Blind defiance regimes

I then consider the set of blind regimes, M^{BD}, by which the sanctioner makes an offer that is unacceptable to both types of target, that is, a *sham offer*. The optimal proposal is determined through a protocol similar to the one followed in the determination of the optimal blind defiance regime, with the difference that, this time, the equilibrium path for both types of target is accept-and-defy. Another difference with the preceding case is the sanctioner's utility, which is now equal to $U(D, D) = haz(k) + (1 - h)\,az(K) - \sigma s$. The sanctioner's utility function records the payoffs generated by the resource denial technology inflicted on the soft and tough types with probabilities h and $1 - h$ respectively at cost σs. As earlier, Lemma 2 is the solution to a constrained maximization.

Lemma 2 (Blind defiance) *The blind defiance regime that maximizes the sanctioner's utility is the menu* $(0 \le t^* \le kZ, 0)$.

The result is intuitive. Target, in this equilibrium, is not delivering the good ($z = 0$). Sanctioner is forced to sanction him, at a cost to herself that she minimizes by setting the sanction to zero ($s = 0$). Although the value of the transfer t is capped so as not to elicit any compliance at all, its actual value otherwise is not pinned down by the equilibrium, because no transfer is ever delivered.

This simple equilibrium yields a utility for the sanctioner that is equal to zero. Since this zero payoff is always better than the war payoff $-w_S$, which is negative, it follows that the blind defiance regime is always preferred to the war outcome. Hence, the following corollary:

Corollary 1 (War) *Subset M^W is dominated by the optimal blind defiance regime.*

6 Revelation regimes

The purpose of the revelation regime is to identify the target type by forcing each type to respond differently to the offer: have soft comply and execute the contract but have tough defy, be sanctioned, be denied resources, and partially comply. To achieve this goal, the sanctioner must find a menu that, in a first step, secures "accept" from both types of target, and then manages to have soft "comply" and tough "defy." Following a now familiar solution protocol, I show in the Appendix that there exists a subset of menus that meet these conditions and yield the sanctioner utility: $U(C, D) = h(aZ - t) + (1-h)(az(K) - \sigma s)$; with probability h the sanctioner faces the soft type with whom she trades the transfer for the good, while with probability $1-h$, she faces the tough type on whom she inflicts a sanction and out of whom she squeezes limited compliance. The sanctioner's problem then is to choose the menu that maximizes her utility among this subset of menus. The results of this calculation are consigned to the following lemma.

Lemma 3 (Revelation) *For subset M^R, the sanctioner's optimal offer is*:

– *the reward only menu* $(kZ, 0)$ *if* $\sigma > \widetilde{\sigma}$,
– *the anything-goes menu* $\left(kZ - w_T \leq t^* \leq kZ,\ 0 \leq s^* \leq \frac{w_T}{1+\zeta}\right.$ *with* $t^* = kZ - \left.(1+\zeta)s^*\right)$ *if* $\sigma = \widetilde{\sigma}$,
– *the mixed menu* $\left(kZ - w_T, \frac{w_T}{1+\zeta}\right)$ *if* $\sigma < \widetilde{\sigma}$ *and* $w_T < kZ$,
– *the sanction only menu* $\left(0, \frac{kZ}{1+\zeta}\right)$ *if* $\sigma < \widetilde{\sigma}$ *and* $w_T \geq kZ$,

with, $\widetilde{\sigma} = \frac{h}{1-h}(1+\zeta) + \frac{a\zeta}{K}$. *In each case, both types of target first accept the offer, while the soft type then complies whereas the tough type defies.*

Unlike blind regimes, revelation regimes use sanctions. Although sanctions are not always imposed ex post—they are not when the interaction reveals the target to be soft—ex ante, however, they are expected to be on the equilibrium path. As a result, resorting to sanction threats is costly in expectation, much like rewards. The sanctioner's choice between sanction and reward, therefore, depends on how they respectively cost her. If the marginal cost of sanctioning σ is high then the sanctioner designs a menu with a reward only. In contrast, if σ is low, then it makes sense for the sanctioner to substitute sanction for reward, keeping in mind that the extent to which she can do so depends on the value of the target's war payoff $-w_T$; she does so fully if $-w_T$ is very low and partly if $-w_T$ is not too low.

The optimal menu, therefore, depends on the value of the two parameters σ and w_T, which I map in Fig. 2, with σ on the horizontal axis and w_T on the vertical axis. As Fig. 2 indicates, there is a cutpoint on the σ axis, denoted $\widetilde{\sigma}$, past which there exists a unique reward-only equilibrium. Before that cutpoint, sanctioning dominates rewarding, fully if w_T is higher than some cutpoint kZ, partly if lower. There is finally the knife-edge case where sanction and reward have identical marginal cost $(\sigma = \widetilde{\sigma})$, yielding an indeterminate result, because any combination of s and t is a solution as long as s is not too low to trigger a war. Being of infinitesimal empirical relevance, this case can be ignored at limited loss in generality.

Fig. 2 Optimal revelation regimes

There is no need to consider the case in which the sanctioner would want to reveal the target's identity so as to extract compliance from the tough type and force defiance on his soft alter ego. Such role reversion is not possible. This result is the object of the next lemma, proven in the Appendix.

Lemma 4 (Role reversal) *There does not exist a revelation regime that allows sanctioner to reward the tough type and sanction the soft type.*

7 When does revelation prevail?

A short answer is: When the threat of war is intermediate—neither too low nor too high—and when the marginal cost of sanctioning is low, holding everything else constant. The complete—and longer—answer is stated in Proposition 1.

Proposition 1 *Revelation regimes dominate all other regimes for values of w_T such that*

– *if* $a > \max(k, K - k)$, $w_T \leq W(\sigma)$ *with*

$$W(\sigma) = \begin{cases} \overline{w_T}(\sigma) \; \textit{if}\, 0 \leq \sigma \leq \min\left(\overline{w_T}^{-1}(kZ), \widetilde{\sigma}\right) \\ \overline{\overline{\overline{w_T}}}(\sigma) \; \textit{if}\, \overline{w_T}^{-1}(kZ) < \sigma \leq \min\left(\overline{\overline{\overline{w_T}}}^{-1}((K-a)Z), \widetilde{\sigma}\right) \\ \overline{\overline{w_T}} \qquad \textit{if}\; \min\left(\overline{\overline{\overline{w_T}}}^{-1}((K-a)Z), \widetilde{\sigma}\right) < \sigma \end{cases}$$

– *if* $k < a < K - k$, $w_T \leq W(\sigma)$ *with*

$$W(\sigma) = \begin{cases} \overline{w_T}(\sigma) \; \textit{if}\, 0 \leq \sigma \leq \min\left(\overline{w_T}^{-1}((K-a)Z), \widetilde{\sigma}\right) \\ \overline{\overline{w_T}} \qquad \textit{if}\; \min\left(\overline{w_T}^{-1}((K-a)Z), \widetilde{\sigma}\right) < \sigma \end{cases}.$$

– *if* $K - k < a < k$, $\underline{w_T}(\sigma) \le w_T \le W(\sigma)$ *over the interval* $0 \le \sigma \le \overline{\overline{\overline{w_T}}}^{-1}((K-a)Z)$ *with* $W(\sigma) = \begin{cases} \overline{w_T}(\sigma) \text{ if } 0 \le \sigma \le \min\left(\overline{w_T}^{-1}(kZ), \tilde{\sigma}\right) \\ \overline{\overline{\overline{w_T}}}(\sigma) \text{ if } \overline{w_T}^{-1}(kZ) < \sigma \le \overline{\overline{\overline{w_T}}}^{-1}((K-a)Z) \end{cases}$

– *if* $a < \min(k, K-k)$, $\underline{w_T}(\sigma) \le w_T \le \overline{w_T}(\sigma)$ *over the interval* $0 \le \sigma \le \overline{w_T}^{-1}((K-a)Z)$, *with* $\overline{w_T}(\sigma) \equiv \left(K - a + ah + \frac{1-h}{1+\zeta}\left(\frac{a\zeta}{K} - \sigma\right)k\right)Z$, $\overline{\overline{w_T}} \equiv (K-a+ah-kh)Z$, $\overline{\overline{\overline{w_T}}}(\sigma) \equiv \frac{(K-a+ah-kh)Z}{(1-h)\left(1-\frac{1}{1+\zeta}\left(\frac{a\zeta}{K}-\sigma\right)\right)}$, *and* $\underline{w_T}(\sigma) \equiv \frac{h(k-a)Z}{h+\frac{1-h}{1+\zeta}\left(\frac{a\zeta}{K}-\sigma\right)}$.

A more direct way of presenting Proposition 1 is to simulate and plot parts of it. The proposition is divided into four clauses. Modeled after Fig. 2 and Fig. 3 maps the results of a simulation for particular parametric values reported below the graph that meet the conditions of the first clause. The weakly decreasing function $W(\sigma)$ delimits the space between an upper region where blind compliance dominates and a lower region where revelation dominates. In contrast, the equilibrium featured in Fig. 4 offers an alternative simulation that corresponds to the third clause of Proposition 1. The space is divided into three regions, an upper region of blind compliance, a lower region of blind defiance, and an intermediate region of revelation in the form of a triangle.

I now provide the intuition behind Proposition 1, first contrasting blind compliance and revelation regimes. We saw earlier that extracting compliance from the target irrespective of his type can be very cheap when sanctions are the main incentive, yet prohibitively costly when transfers have to be used along with sanctions. In contrast, the cost of a revelation regime is moderate—sanctions and transfers are calculated to be just enough to deter the soft type from defying and to entice that same type to comply. This means that in relation to blind compliance, revelation is suitable in two situations: when transfers have to be used (because w_T is low) and when the marginal cost of sanctioning, σ, is low.

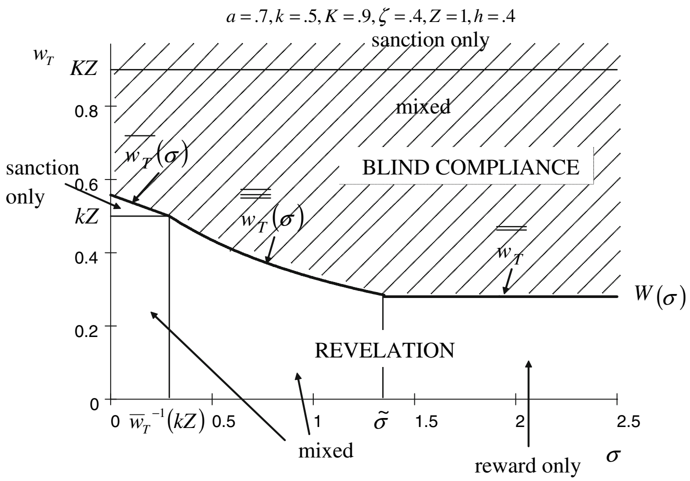

Fig. 3 Simulated equilibrium for $a > \max(k, K-k)$

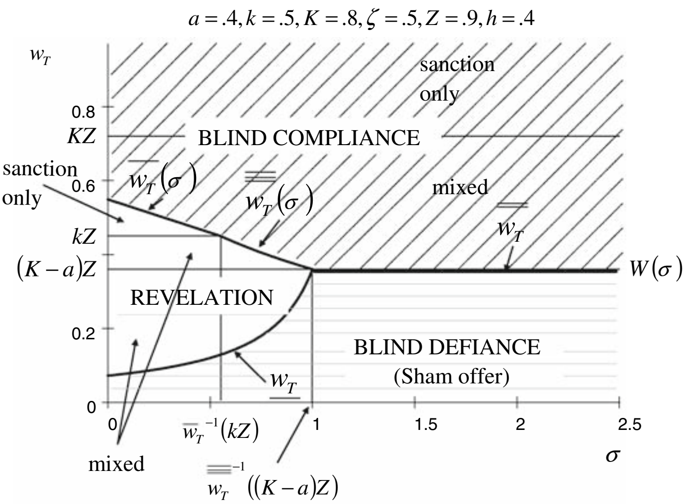

Fig. 4 Simulated equilibrium for $K - k < a < k$

I then contrast blind defiance and revelation regimes in Fig. 4. Two remarks are important. First, blind defiance occurs in Fig. 4 and not in Fig. 3. This is because in Fig. 3, revelation dominates blind defiance in all cases, whereas in Fig. 4, revelation dominates blind defiance in certain parts of the space. The decisive factor here is how much sanctioner and soft target desire good Z. The sanctioner's desire is captured by marginal gain a, while the soft target's resolve is captured by marginal cost k. In the revelation regime, the equilibrium transfer is calculated to match the soft target's cost of compliance kZ (assume for now that the sanction component s is very small). If that transfer is lower than what the sanctioner gains from the transaction (payoff aZ), that is, if $kZ < aZ$ or $k < a$, then the sanctioner prefers making the real to the sham offer—this is the case depicted in Fig. 3. In contrast, if the transfer is higher than what the sanctioner gains from the transaction, i.e., if $a < k$, then the sanctioner is better off making the sham offer and get zero instead of a negative payoff. In this second case, which is simulated in Fig. 4, the sham offer takes over the parametric range that the revelation regime would have occupied, but not fully: there still is a niche for revelation for somewhat higher values of s (and thus lower values of t) and lower values of σ.

8 Comparative statics

The comparative statics presented here bears only on the compliance structure of the sanction regime—I leave aside the incentive structure.[10] Comparative statics indicates that the revelation regime dominates when the sanctioner's marginal cost of sanctioning σ is low, the resource denial technology is efficient, and target is likely to be the soft

[10] As it can be guessed from Figs. 3 and 4, the incentive structure is continuous within each regime, but discontinuous across regimes.

A rise in parameter …		leads to a change in the range of the mechanism with sign …		
		Blind Defiance	revelation	blind compliance
Sanctioner's marginal cost of sanctioning	σ	+	-	+
Efficiency of resource denial technology	ζ	-	+	-
Type Differential	K-k	+*	-/+*	-*
Sanctioner's marginal gain of compliance	a	-	+/-	+
Likelihood target is soft	h	-	+	-
Target's Cost of war	w_T	-	+/-	+

+/- means that revelation gains from blind defiance but loses to blind compliance.
-/+ means that revelation loses to blind defiance but gains from blind compliance.
* Conditioned on *h* taking an intermediate value.

Fig. 5 Comparative statics

type. The net impact of a rise in any of the other parameters—uncertainty, sanctioner's marginal gain of compliance and target's cost for war—is a priori indeterminate. Blind compliance regimes, in contrast, tend to dominate when the sanctioner's two marginals are high and the target's cost of war is high. Blind defiance tend to dominate when the sanctioner's marginal cost of compliance is high and uncertainty is high. Figure 5 summarizes the main comparative statics results. I develop the intuition behind each one; proofs can be found in the Appendix.

A rise in the sanctioner's marginal cost of sanctioning, σ, makes revelation less efficient than blind compliance, and, in the case where the sanctioner is less resolute than the soft target (which is the context envisaged in Fig. 4), less efficient also than blind defiance. This is because the revelation regime is the only one to require in expectation the execution of the sanction threat. These results can be directly read off Figs. 3 and 4.

A related finding is that revelation regimes tend to be preferred when the resource denial technology is more efficient, that is, when ζ is close to 1. A more efficient resource denial technology increases the efficiency of actual sanctioning.

Revelation regimes are preferred to blind compliance regimes but lose to blind defiance regimes in situations where uncertainty has a big impact on sanctioner's payoff, i.e., when the type differential, $K - k$, is large. It is easy to see that if the two types, though dissimilar, were not that far apart from each other, say $K = k + \varepsilon$ with $\varepsilon > 0$ and very small, the lack of common knowledge would be of limited practical relevance. For instance, it is easy to verify that the transfer that it would take to get soft to comply in a revelation regime would not be that different from what it would take to get tough, and a fortiori soft, to comply in a blind compliance regime, making the two categories almost interchangeable. In contrast, a large ε makes information asymmetry an important feature of the game. Now, what happens when the type differential increases? One way of finding out is to calculate the impact of an increase in K while holding k constant. This is performed in the Appendix.[11] Two effects are noticeable. A first effect is that revelation encroaches on blind compliance. The reason for this result is that blind compliance regimes waste incentives on the soft target type, who is willing to comply for less than what is threatened or promised. Revelation regimes, in contrast, allow the sanctioner to behave like a monopolist; she segments her market and differentiates her product, rewarding the soft type at the marginal cost while sanctioning the tough type. A second effect is that revelation in turn is displaced by blind defiance whenever the parametric conditions for blind defiance are met (the context of Fig. 4). The reason is that an increase in the tough type's marginal cost of compliance, K, raises the cost of revelation in the form of a higher sanction cost, making the sanctioner better off giving up altogether.

A sanctioner with a high marginal value for compliance, a, will prefer blind compliance to revelation and revelation to blind defiance. The blind compliance regime delivers full compliance from both types, whereas the revelation regime, in contrast, delivers full compliance from the soft type but partial compliance only from the tough type. The blind defiance regime, in contrast, delivers no compliance at all. This comparative statics is conditioned on the frequency of soft types, h, having an intermediate value; too many Softs make revelation similar to, though less expensive than, blind compliance, whereas not enough of them price out any form of compliance.

A rise in the frequency of soft types, h, in relation to tough types, $1-h$, increases the relevance of revelation regimes vis-a-vis all forms of blind regimes. Revelation is most efficient when facing a soft type, because the reward for compliance is set at Soft's marginal value, kZ. Revelation is less efficient when facing a tough type, because it requires the imposition of a sanction, a sheer loss for sanctioner. Increasing the proportion of soft types increases the chance of extracting compliance at Soft's marginal cost while decreasing the risk of imposing the sanction, thereby making revelation more efficient. Simultaneously, a rise in the proportion of soft types makes blind compliance regimes less efficient, because blind compliance wastes resources on the soft type.

Finally, a rise in the target's cost for war, w_T, makes blind compliance more efficient all around because it reduces the price that sanctioner pays in the mixed menu, $KZ-w_T$. Although the very same effect is present in the mixed menu of the revelation regime, this positive effect is offset by a negative one in the form of a rise in the cost of the imposed sanction, $\frac{w_T}{1+\zeta}$. This result can directly be read off Figs. 3 and 4.

I now offer three paradigmatic cases, one for each sanction regime. The purpose of these sections is not to prove, but illustrate.

9 An example of revelation: engagement toward Iraq (1989)

A first illustration of the model's predictions is the policy of engagement toward Iraq laid out in National Security Directive 26 and pursued by the Bush senior administra-

[11] Another way is to allow k to drop while holding K constant. Both ways are calculated in the Appendix; they yield comparable results.

tion from 1989 until August 1990. I first present the story and then discuss the fit with the comparative statics derived in the prior section.

Following the Iran-Iraq War, the Bush-senior Administration in its NSD 26 recognized the need, according to Kenneth Juster, a senior official then at the State Department, to "probe and test the Iraqis as part of the effort to moderate their behavior" (Juster 2000, p. 56). Washington confronted Saddam with two options: either give up any illegal use of chemical or biological weapons, respect international nuclear safeguards, and settle with Iran and be rewarded with trade and aid; or else face the continuation or worsening of economic and diplomatic sanctions. Brent Scowcroft, who occupied the position of national security adviser, summarized the policy as a "mix of limited incentives and strong disincentives."[12] Therefore, NSD 26 suggested that the United States face Saddam with a revelation regime mixing carrots and sticks.

The revelatory nature of the regime is the reason why the policy met broad support within the Arab world, among Europeans, virtually every American expert on the Middle East, and Congress. In essence, Washington de facto offered Saddam a deal that was considered as acceptable by all. When he failed to deliver, by assaulting Kuwait, he revealed to the international community a brashness that made it impossible for Iraq's traditional allies to come to his support. Such a consensus reflected the fact that the United States had first tried to engage Saddam and, so doing, had helped reveal Saddam's true ambitions to the world.

How does the choice of a revelation regime fit with the comparative statics? I develop each of the six parameters in the order listed in Fig. 5. First, the cost of sanctioning Iraq for the United States was rather low. Iraq was already under a regime of sanction in the wake of its war with Iran. Further sanctions in the form of an export ban on nuclear and military technology would have no consequences on U.S. industry. Even a boycott of Iraqi oil exports would have but a limited impact on world prices. Second, the resource denial technology was thought to be efficient. Nuclear weapon countries at the time believed that the NPT regime, which consisted of export controls imposed by countries belonging to the Nuclear Suppliers Group (NSG) and monitoring conducted by the International Atomic Energy Agency (IAEA), was efficient in curbing proliferation.

Third, the type differential was high. Given the opacity of his rule, both dictatorial and secretive, Saddam was considered as an unknown quantity in Washington. Not only was there a good chance that Saddam's fiery rhetoric was not reducible to mere posturing for leverage, but also that he could be a very dangerous type. There was no obvious limited territorial goal that an Iraqi leader feeling chronically insecure could strive to meet that would satisfy him—$K - k$ was probably high.

Fourth, the United States had a stake in turning Iraq into a country with which it could do business. Washington was especially interested in curbing Iraq's appetite for nuclear weapons. The stakes, however, were limited in that Iraq was no direct threat to U.S. security—a was moderately high.

Fifth, those in Washington who were behind NSD 26 believed that Iraq was ready to change course. Saddam could be made to work with the United States. There was

[12] *The Washington Post* October 10, 1992, Saturday, Final Edition.

a widespread belief among State Department officials in light of an ongoing aggiornamento taking place in Bagdad (weakened relations with Moscow, non-aggression pact with Saudi Arabia) that Saddam was open to engagement—h was high.[13] Last, Iraq's expected cost of war were it to declare war on the United States or of its allies in the region was probably high in light of the power imbalance.

In sum, the engagement policy was compatible with the parametric alignment that is conducive to revelation.

10 An example of blind compliance: South Korea and nuclear proliferation (1976)

The second case study illustrates the blind compliance regime. South Korea signed the Nuclear Proliferation Treaty in 1968 and ratified it in 1975, probably under U.S. pressure. Yet, months after ratification, the South Koreans announced their intention to buy a reprocessing facility from France, allowing it to separate plutonium from spent fuel, arguably to reuse it in power generation, but more likely, as the U.S. believed, to use in a nuclear bomb. Washington reacted by mixing the threat of blocking further delivery of fuel and equipment to a U.S.-made reactor under construction with the promise of not withdrawing extant American troops. According to one observer, "(T)he South Korean government was undoubtedly most influenced by the fear that its own actions were endangering the vital U.S. security connection" (Yager 1985, p. 190). As a Korean diplomat said, "the United States made the strongest possible representations to the Korean and French governments."[14]

The sanction regime geared to South Korea is a fit illustration of a blind compliance regime. The choice of blind compliance, I argue, reflected many of the predicted parametric values, starting with the U.S.'s marginal costs of sanctioning and compliance, which were closely linked in this case. The cost of sanctioning, and conversely, the gain from compliance, were huge, not so much in military than in reputational terms. In the midst of the Cold War, the United States could not turn against South Korea, a country literally located on the frontline. With Congress intent on reducing American troops abroad, it was in the interest of the executive to hush up the crisis with Seoul as quietly and quickly as possible by giving in to their demand for a steady U.S. military presence in South Korea, notwithstanding the fact that Washington had at its disposal a very efficient resource denial technology—embargo of nuclear fuel—a resource which, in other circumstances, could have supported a successful strategy of revelation.

The type differential, next, was small. Despite the repressive nature of Korean President Park Chung Hee's rule, the variance of the possible strategic views that he could hold, in light of the history of the Korean peninsula, was limited. The security pressure for South Korea was so strong and focused that he would have found it difficult to push for anything other than a credible demonstration of the American security guarantee; $K - k$ was low.

[13] The point is made by Juster (2000).

[14] *The New York Times*, 1 February 1976, p. 11.

Given the intensity of the Korean concern for security, it was very likely that Park was tough; he meant to pursue a secret nuclear weapons program if the United States did not provide him with sufficient guarantees. Last, the cost of a war with the United States would be devastating for South Korea (high w_T).

Therefore, with the exception of the resource denial technology, which was high, the parametric line-up in this case favored blind compliance.

11 An example of blind defiance: Italy's invasion of Abyssinia (1935)

According to the model, a sham offer is one with a zero sanction threat and an insufficient transfer. This is a much too restrictive definition, one that reflects the parsimony of the model, leaving out sanctioning rationales other than compliance and information revelation, such as retribution, reputation building, and the protection of domestic interests. The presence of these other rationales in reality requires the present definition of a sham offer to be extended to any offer that is calculated not to elicit compliance from any type. A most notable case of sham offer is the 1935 League of Nations' sanction of Italy in response to the latter's unprovoked invasion of Abyssinia.

Italy being allied with Britain and France in their competition with Nazi Germany, these two countries made sure that Italy would be punished just enough to uphold the principles of the League, but not to the point of alienating Mussolini and forcing him to side with Hitler. Hence, the member states embargoed the supply of arms and strategic materials and agreed to cut loans and credits to the Italian government, but under French and British pressure, they never embargoed oil, even though such an embargo would have literally grounded the Italian intervention force.[15] Eventually, Italian troops entered Addis Ababa and soon after the League discontinued sanctions.

The main reason for the timidity of the League's action was the high risk of war and consolidation of the "axis." Mussolini explicitly told British Foreign Minister Eden that he would look upon an oil embargo as an instance of *casus belli* (Hoffmann 1967, p. 144). Combined with a high cost of sanctioning, the British and the French had a low preference for compliance. The marginal benefit of denying Abyssinia to the Italians was low relative to the marginal cost of imposing the sanction, with Rome threatening to repudiate its foreign debt. Hence, a was low while σ was high. Although there existed an efficient resource denial technology in the form of oil embargo, its use was impractical in the presence of a high risk of war. The sanctions that were eventually imposed were perfunctory (low ζ).

There was a strong probability that Mussolini was tough (h low), afterall virility was il Duce's declared motto, yet it is unclear how much wide the type differential actually was. In the worst case, Mussolini would take over Ethiopia, one of the last African nations that was not a British or French colony[16]—Mussolini was no Hitler ($K - k$ was low). Finally, the expected Italian cost of an eventual war with Britain or France, though not low, was much lower than in the two previous case studies,

[15] See case 35-1 in Hufbauer et al. (1990, vol. 2).

[16] In the best case, the invasion would see a repeat the defeat of the First Italo-Abyssinian War of 1896.

especially if we consider the possibility of Germany joining in on the Italian side (w_T was relatively low).

In sum, with the exception of the resource denial technology, the other parameters take values that match those predicted by the model in a case of blind defiance.

12 Conclusion

The theory presented here offers an alternative to signaling theories of sanctioning in the face of uncertainty. It gives the sanctioner a choice between pursuing a strategy of information revelation, blind compliance, or mere grandstanding through a sham offer. Revelation regimes can be used in a first step to determine the degree of resolve of the target and in a second step to reward the more compliant type and sanction the more resilient type. Blind regimes aim at eliciting compliance irrespective of resolve. Sham offers serve as useful umbrella for sanctioning strategies that are motivated by goals that fall short of compliance or information revelation. Within each category of regime, reward and punishment are calibrated so as to achieve the desired goal. Rewards are preferred over sanctions if the cost of war is high and, in the case of revelation regimes, if the cost of sanctioning is high.

The article illustrates the argument by offering three sanction cases selected for their fit with the predictions of the model. The sanctioner prefers a revelation regime to a blind regime in cases of high cost of war or uncertainty. In contrast, a blind regime is preferred when the sanctioner has a strong preference for compliance or in the presence of low uncertainty. Sham offers tend to be preferred when the sanctioner's marginal gain for compliance is low and its marginal cost of sanctioning is high.

The revelation regime is useful to a sanctioner who needs to know a target's resolve. It is also useful to a sanctioner who already knows, but needs to convince a domestic or an international audience. In the post-Cold War era, where effective action is collective action, acquiring intelligence is necessary but insufficient. Complete information is a prerequisite for collective action, and sanction regime design is the sure way of eliciting complete information.

The revelation regime should be of special interest to students of norms, for if the purpose of a norm is to draw a line between normative types who abide by the norm and so-called "rogues" who don't, then it is important that types be revealed precisely and easily (see Nossal 1989).

Appendix: Proofs

Proof of Lemma 1

The sanctioner who wants to maximize her objective function by means of a blind compliance regime faces the following problem:

$$\max_{t,s} U\left(C, C\right) = aZ - t \qquad \text{(A1)}$$

subject to several constraints that insure that the target goes down the blind compliance equilibrium path. A first constraint is that the tough target's utility for compliance be equal or better than his payoff for war, $-w_T$; formally, $t - KZ \geq -w_T$. The same is true for soft, $t - kZ \geq -w_T$. Another constraint must insure that both types of target comply rather than incur the sanction. Formally, $t - KZ \geq -s - Kz(K)$. The same is true for soft, $t - kZ \geq -s - kz(k)$. To complete the setup of the sanctioner's maximization problem, one must add three more constraints: $s \geq 0$ to insure that sanction s does not turn into an incentive; $t \geq 0$ to insure that transfer t does not turn into a sanction; and $z(\kappa) = \frac{\zeta}{\kappa}s$, with $\kappa \in \{k, K\}$, the resource denial technology. After substituting the latter into the other constraints and discarding the second and fourth constraints on the grounds that the others make them redundant, the set of constraints can be rewritten as

$$t - KZ \geq -w_T, \tag{A2}$$

$$t - KZ \geq -(1+\zeta)s, \tag{A3}$$

$$s \geq 0, \tag{A4}$$

$$t \geq 0. \tag{A5}$$

I use λ, γ, ϕ and ψ, all positive or naught, as multipliers for the four constraints in the order listed. The objective function is semi-concave. The Kuhn-Tucker conditions are twofold: a non-binding constraint implies that its corresponding parameter is zero (KT1); a positive parameter implies that the constraint is binding (KT2) (Kreps 1990, p. 777). The maximization yields two first-order conditions with respect to s and t successively

$$(1+\zeta)\gamma + \phi = 0, \tag{A6}$$

$$\lambda + \gamma + \psi = 1. \tag{A7}$$

From the first order condition, it follows that $\gamma, \phi = 0$. This implies that constraints (A3) and (A4) are not necessarily binding and that the equilibrium value of the sanction is indeterminate within the range $\begin{cases} \frac{w_T}{1+\zeta} \leq s^* & \text{if } t^* = KZ - w_T \\ \frac{KZ}{1+\zeta} \leq s^* & \text{if } t^* = 0 \end{cases}$. The lower bounds are derived by plugging into (A3) the equilibrium values of t calculated below. Indeterminacy follows from the fact that sanctioning in a blind compliance menu is located off the equilibrium path.

With respect to λ and ψ, three cases must be considered:

(1) If $t^* \geq KZ$, (A2), (A5), and KT1 imply that $\lambda, \psi = 0$, in contradiction with (A7).

(2) If $0 < t^* < KZ$, (A5), KT1 and (A7) imply that $\psi = 0$ and $\lambda > 0$. By KT2, (A2) yields $t^* = KZ - w_T$, a mixed menu for $w_T \in (0, KZ)$.

(3) If $t^* = 0$, as seen earlier, s is indeterminate within the parametric range $s^* \geq \frac{KZ}{1+\zeta}$, for $w_T \geq KZ$. Combining the results in 2 and 3 yields Proposition 1.

Proof of Lemma 2

The sanctioner who wants to maximize her objective function by means of a blind defiance regime solves the following problem:

$$\max_{s,t} U(D, D) = h(az(k) - \sigma s) + (1-h)(az(K) - \sigma s) \tag{A8}$$

subject to the two peace constraints $-s - kz(k) \geq -w_T$ and $-s - Kz(K) \geq -w_T$, the two incentive constraints $-s - kz(k) \geq t - kZ$ and $-s - Kz(K) \geq t - KZ$, the two boundary constraints $s, t \geq 0$, and $z(\kappa) = \frac{\zeta}{\kappa}s$, with $\kappa \in (k, K)$, the resource denial technology. After substituting the latter into the other constraints and discarding the first and fourth constraints on the grounds that the others make them redundant, one is left with four constraints. The objective function is semi-concave. Assume λ, θ, ψ and ϕ, with $\lambda, \theta, \psi, \phi \geq 0$, are the multipliers for the four remaining constraints in the order listed. The maximization yields two first-order conditions with respect to s and t successively.

$$-\lambda(1-\zeta) - \theta(1-\zeta) + \psi = h\left(\sigma - \frac{a\zeta}{k}\right) + (1-h)\left(\sigma - \frac{a\zeta}{K}\right), \tag{A9}$$

$$-\theta + \phi = 0. \tag{A10}$$

By (A9), $\psi > 0$, given that the right-hand side of (A9) is positive—this follows from the assumption that $\sigma > a\left(h\frac{\zeta}{k} + (1-h)\frac{\zeta}{K}\right)$. This means that $s = 0$, $z = 0$. The value of t is left indeterminate within the interval $(0, kZ)$ by the third constraint, for off the equilibrium path. Formally, if $\theta > 0$, by (A10) $\phi > 0$, and $t = kZ$. But if $\theta = 0$ and thus $t \leq kZ$, then by (A10) $\phi = 0$ and $t \geq 0$—the value of t cannot be pinned down.

Proof of Lemma 3

The sanctioner solves problem $\max_{s,t} U(C, D) = h(aZ - t) + (1-h)(az(K) - \sigma s)$, subject to two peace constraints $-s - Kz(K) \geq -w_T$ and $t - kZ \geq -w_T$, the first for the tough target type, the second for the soft, and two incentive constraints, $-s - Kz(K) \geq t - KZ$ and $t - kZ \geq -s - kz(k)$ for the tough and soft types respectively. Furthermore, as for blind regimes, two more constraints are necessary: $s \geq 0$, $t \geq 0$. Last, $z(\kappa) = \frac{\zeta}{\kappa}s$, with $\kappa \in \{k, K\}$. Substitute for the value of $z(\kappa)$ in the maximization equation and the other constraints, and drop the second peace constraint for being redundant in light of the first and last. The sanctioner's problem can be restated as:

$$\max_{s,t} U(C, D) = h(aZ - t) + (1-h)\left(\frac{a\zeta}{K} - \sigma\right)s \tag{A11}$$

such that

$$-(1+\zeta)\,s \geq -w_T, \tag{A12}$$
$$-(1+\zeta)\,s \geq t - KZ, \tag{A13}$$
$$t - kZ \geq -(1+\zeta)\,s, \tag{A14}$$
$$s \geq 0, \tag{A15}$$
$$t \geq 0. \tag{A16}$$

I use λ, ν, μ, ϕ and ψ, with $\lambda, \nu, \mu, \phi, \psi \geq 0$, as multipliers for the five constraints in the order I just listed them. The maximization yields the first-order condition with respect to s

$$-(\lambda+\nu-\mu)(1+\zeta)+\phi = (1-h)\left(\sigma - \frac{a\zeta}{K}\right), \tag{A17}$$

Note that the right hand side of (A17) is positive because of the assumption that $\sigma > a\left(h\frac{\zeta}{k} + (1-h)\frac{\zeta}{K}\right)$. The first-order condition with respect to t is

$$-\nu + \mu + \psi = h. \tag{A18}$$

There are five cases in which the five multipliers can either take a positive or zero value. To show this, note that one cannot have $\nu > 0$ and $\mu > 0$ simultaneously, for, if we did, (A13) and (A14) would imply after subtraction of the two equations that $Z < z$, a contradiction. For now, assume that $\nu = 0$, $\mu > 0$, and thus $t - kZ = -s(1+\zeta)$. Then consider the following cases:

(1) If $t = kZ$, then $t > 0$ and $s = 0$, implying $\phi \geq 0$ and $\lambda = \psi = 0$. Substituting these values into the first-order conditions and extracting σ yields $\sigma = \frac{h(1+\zeta)+\phi}{1-h} + \frac{a\zeta}{K}$ and thus $\sigma > \frac{h}{1-h}(1+\zeta) + \frac{a\zeta}{K} \equiv \widetilde{\sigma}$ if $\phi > 0$. $\phi = 0$ is the limit case where the solution $s = 0$ is optimal at the corner ($\sigma = \widetilde{\sigma}$).

(2) If $0 < t < kZ$, then $s > 0$, implying that $\phi = \psi = 0$. Substituting these values along with $\nu = 0$ and $\mu > 0$ into the first-order conditions yields $\lambda(1+\zeta) = h(1+\zeta) - (1-h)\left(\sigma - \frac{a\zeta}{K}\right)$, thus implying two possibilities. If $\sigma = \widetilde{\sigma}$, then $\lambda = 0$, and, by (A12), s can take any value within the $\left(0, \frac{w_T}{1+\zeta}\right)$ interval while $t^* = kZ + (1+\zeta)\,s^*$;

(3) whereas if $\sigma < \widetilde{\sigma}$, then $\lambda > 0$, $s^* = \frac{w_T}{1+\zeta}$, and $t^* = kZ - w_T$.

(4) If $t^* = 0$, by (A12) $s^* = \frac{kZ}{1+\zeta}$. This implies that $\phi = 0$, $\lambda \geq 0$, and $\psi \geq 0$. If $\lambda = 0$, $s^* < \frac{w_T}{1+\zeta}$. This case obtains when $w_T > kZ$ and $\sigma = \frac{(h-\psi)(1+\zeta)}{1-h} + \frac{a\zeta}{K}$, and thus $\sigma < \widetilde{\sigma}$ if $\psi > 0$ or $\sigma = \widetilde{\sigma}$ if $\psi = 0$.

(5) While if $\lambda > 0$, $s^* = \frac{w_T}{1+\zeta}$ and so $w_T = kZ$ and $\sigma = \frac{(h-\psi-\lambda)(1-g)(1+\zeta)}{1-h} + \frac{a\zeta}{K}$, which is always inferior to $\widetilde{\sigma}$ irrespective of the value of ψ.

Then, consider the case where $\nu > 0$ and $\mu = 0$: It follows that $t - KZ = -(1+\zeta)\,s$ and, from (A18), that $\psi > 0$, $t^* = 0$. Combining these two results yields $s^* = \frac{KZ}{1+\zeta}$. From (A17), however, $\phi > 0$, implying $s^* = 0$, a contradiction.

Last, consider the case where $\nu, \mu = 0$: From (A17), $\phi > 0$ and thus $s^* = 0$. From (A12), we have $t \geq kZ$ and thus $\psi = 0$, yielding by (A18) that $h = 0$, a contradiction.

Proof of Lemma 4

A revelation regime that would lead soft to defy and be punished while leading tough to comply and be rewarded would be equivalent to the following sanctioner's problem: $\max U(D, C) = h(az(k) - \sigma s) + (1-h)(aZ - t)$, subject to the two peace constraints $t - KZ \geq -w_T$ and $-s - kz(k) \geq -w_T$, the first for the tough target type, the second for the soft, and two incentive constraints, $t - KZ \geq -s - Kz(K)$ and $-s - kz(k) \geq t - kZ$ for the tough and soft types respectively. Two more constraints are necessary: $s \geq 0$, $t \geq 0$. Last, $z(\kappa) = \frac{\zeta}{\kappa}s$, with $\kappa \in \{k, K\}$.

However, given that $Kz(K) = kz(k) = \zeta s$, the two incentive constraints are incompatible. Strung together, they yield $t - KZ \geq -(1+\zeta)s \geq t - kZ$, implying that $K \leq k$, a contradiction.

Proof of Proposition 1

Several cases need to be considered along several dimensions: w_T, σ, a, and k. For values of $w_T \geq KZ$, blind compliance dominates. Indeed, the sanctioner's payoff for blind compliance is equal to aZ; in contrast, revelation yields

$$U(C, D) = \begin{cases} \left(ha + \frac{1-h}{1+\zeta}\left(\frac{a\zeta}{K} - \sigma\right)k\right)Z & \text{if } \sigma < \widetilde{\sigma}; \\ h(a-k)Z & \text{if } \sigma > \widetilde{\sigma}; \end{cases}$$

both results are less than aZ. Blind defiance yields payoff 0; it is also dominated.

For values of w_T inferior to KZ, two separate cases must be considered:

(1) First, assume that $a > k$ and $a > K - k$ (and thus $kZ > (K-a)Z$). Consider the following cases:

(a) $kZ \leq w_T < KZ$; blind compliance yields $aZ - KZ + w_T$, which is positive,[17] thereby dominating blind defiance. Revelation yields the same two cases as above, making revelation preferred to blind compliance whenever $\sigma < \widetilde{\sigma}$ and $w_T < \left(K - a + ah + \frac{1-h}{1+\zeta}\left(\frac{a\zeta}{K} - \sigma\right)k\right)Z \equiv \overline{w_T}(\sigma)$, and whenever $\sigma > \widetilde{\sigma}$ and $w_T < (K - a + ah - kh)Z \equiv \overline{\overline{w_T}}$.

(b) $\max(0, (K-a)Z) \leq w_T < kZ$; this case is similar to the prior one if $\sigma > \widetilde{\sigma}$. It differs if $\sigma < \widetilde{\sigma}$, for revelation yields $h(aZ - kZ + w_T) + \frac{1-h}{1+\zeta}\left(\frac{a\zeta}{K} - \sigma\right)w_T$, making revelation preferred whenever $w_T < \frac{(K-a+ah-kh)Z}{(1-h)\left(1-\frac{1}{1+\zeta}\left(\frac{a\zeta}{K}-\sigma\right)\right)} \equiv \overline{\overline{\overline{w_T}}}(\sigma)$. Note

[17] From $a > K - k$, derive $aZ - KZ > -kZ$. From $w_T \geq kZ$ derive $-kZ \geq -w_T$. Bring the two results together to get $aZ - KZ > -w_T$, implying the result $aZ - KZ + w_T > 0$.

that the denominator is always positive for $\sigma \in (0, \widetilde{\sigma})$, while the numerator is positive. In either case, blind defiance yields zero; it is dominated.

(c) $0 < w_T < (K - a)\, Z$; the blind compliance menu now yields a payoff inferior to zero. It is thus dominated by the blind defiance menu, yielding 0. The optimal revelation menu in turn yields

$$U\,(C, D) = \begin{cases} h\,(aZ - kZ + w_T) + \frac{1-h}{1+\zeta}\left(\frac{a\zeta}{K} - \sigma\right) w_T & \text{if } \sigma < \widetilde{\sigma} \\ h\,(a-k)\,Z & \text{if } \sigma > \widetilde{\sigma} \end{cases}.$$

In the first case, revelation is preferred whenever $w_T > \frac{h(k-a)Z}{h + \frac{1-h}{1+\zeta}\left(\frac{a\zeta}{K} - \sigma\right)} \equiv \underline{w_T}\,(\sigma)$. This is always true since, given that $a > k$, the numerator is negative and the denominator is always positive for $\sigma \in (0, \widetilde{\sigma})$. In the second case, revelation is always preferred to blind defiance once again because $a > k$.

Note the following properties of the boundary functions so far identified: $\overline{w_T}^{-1}(kZ) = \overline{\overline{w_T}}^{-1}(kZ)$; $\overline{w_T}\,(\widetilde{\sigma}) = \overline{\overline{w_T}}\,(\widetilde{\sigma}) = \overline{\overline{w_T}}$, $\partial \overline{w_T}\,(\sigma)/\partial\sigma < 0$, and $\partial \overline{\overline{w_T}}\,(\sigma)/\partial\sigma < 0$. In sum, the frontier between blind compliance and revelation regimes in the (σ, w_T) space is a continuously declining function $W_1\,(\sigma)$ as σ moves from 0 to some cutpoint, at which point it stabilizes and turns into a constant.

Putting all that precedes together yields that revelation regimes dominate all other regimes for values of w_T such that $w_T \le W\,(\sigma)$ with

$$W\,(\sigma) = \begin{cases} \overline{w_T}\,(\sigma) \text{ if } 0 \le \sigma \le \min\left(\overline{w_T}^{-1}(kZ), \widetilde{\sigma}\right) \\ \overline{\overline{w_T}}\,(\sigma) \text{ if } \overline{w_T}^{-1}(kZ) < \sigma \le \min\left(\overline{\overline{w_T}}^{-1}((K-a)\,Z), \widetilde{\sigma}\right) \\ \overline{\overline{w_T}} \quad \text{ if } \min\left(\overline{\overline{w_T}}^{-1}((K-a)\,Z), \widetilde{\sigma}\right) < \sigma \end{cases}.$$

The three remaining cases are analyzed by following the same steps, which I omit.

(2) Now assume that $a > k$ and $a < K - k$; revelation regimes dominate for values of w_T such that $w_T \le W\,(\sigma)$ with

$$W\,(\sigma) = \begin{cases} \overline{w_T}\,(\sigma) \text{ if } 0 \le \sigma \le \min\left(\overline{w_T}^{-1}((K-a)\,Z), \widetilde{\sigma}\right) \\ \overline{\overline{w_T}} \quad \text{ if } \min\left(\overline{w_T}^{-1}((K-a)\,Z), \widetilde{\sigma}\right) < \sigma \end{cases}.$$

(3) Assume $K - k < a < k$; the results are the same as in (1) except for case (c) in which expression $\underline{w_T}\,(\sigma)$ is always positive and $\lim_{\sigma \to \widetilde{\sigma}} \underline{w_T}\,(\sigma) = \infty$. Revelation regimes dominate for values of w_T such that $\underline{w_T}\,(\sigma) \le w_T \le W\,(\sigma)$ over the interval $0 \le \sigma \le \overline{\overline{w_T}}^{-1}((K-a)\,Z)$ with $W\,(\sigma) = \begin{cases} \overline{w_T}\,(\sigma) \text{ if } 0 \le \sigma \le \min\left(\overline{w_T}^{-1}(kZ), \widetilde{\sigma}\right) \\ \overline{\overline{w_T}}\,(\sigma) \text{ if } \overline{w_T}^{-1}(kZ) < \sigma \le \overline{\overline{w_T}}^{-1}((K-a)\,Z) \end{cases}$.

(4) Last assume that $a < k$ and $a < K - k$; revelation regimes dominate for values of w_T such that $\underline{w_T}\,(\sigma) \le w_T \le \overline{w_T}\,(\sigma)$ over the interval $0 \le \sigma \le \overline{w_T}^{-1}((K-a)\,Z)$.

Comparative statics

σ : *sanctioner's marginal cost of sanctioning*

$\frac{\partial \overline{w_T}(\sigma)}{\partial \sigma} < 0$, $\frac{\partial \overline{\overline{w_T}}}{\partial \sigma} = 0$, $\frac{\partial \overline{\overline{\overline{w_T}}}(\sigma)}{\partial \sigma} < 0$, $\frac{\partial \underline{w_T}(\sigma)}{\partial \sigma} > 0$ in the relevant $k > a$ range. The $\overline{w_T}(\sigma)$ and $\overline{\overline{\overline{w_T}}}(\sigma)$ functions are decreasing, causing a rise in σ to reduce the corresponding value of w_T and thus the range of such values for which revelation obtains. The $\underline{w_T}(\sigma)$ function is increasing, leading a rise in σ to reduce the range of w_T values for which revelation obtains.

ζ : *resource denial technology*

$\frac{\partial \overline{w_T}(\sigma)}{\partial \zeta} > 0$, $\frac{\partial \overline{\overline{w_T}}}{\partial \zeta} = 0$, $\frac{\partial \overline{\overline{\overline{w_T}}}(\sigma)}{\partial \zeta} > 0$, $\frac{\partial \underline{w_T}(\sigma)}{\partial \zeta} < 0$, $\frac{\partial \widetilde{\sigma}}{\partial \zeta} > 0$, $\frac{\partial}{\partial \zeta} \overline{w_T}((K-a)Z)^{-1} > 0$, $\frac{\partial}{\partial \zeta} \overline{w_T}(kZ)^{-1} > 0$, $\frac{\partial}{\partial \zeta} \overline{\overline{\overline{w_T}}}((K-a)Z)^{-1} < 0$; a rise in ζ shifts the $W(\sigma)$ schedule up and to the right, while it lowers the $\underline{w_T}(\sigma)$ function, opening up space for the revelation regime.

K : *tough target type's marginal cost of compliance*

$\frac{\partial \overline{w_T}(\sigma)}{\partial K} > 0$ if $a < \frac{(1+\zeta)K^2}{(1-h)\zeta k}$, negative otherwise; $\frac{\partial \overline{\overline{w_T}}}{\partial K} > 0$, $\frac{\partial \overline{\overline{\overline{w_T}}}(\sigma)}{\partial K} > 0$, $\frac{\partial \underline{w_T}(\sigma)}{\partial K} > 0$ if $a < k$, the domain on which $\underline{w_T}(\sigma)$ is defined; $\frac{\partial \widetilde{\sigma}}{\partial K} < 0$, $\frac{\partial}{\partial K} \overline{w_T}((K-a)Z)^{-1} < 0$, $\frac{\partial}{\partial K} \overline{w_T}(kZ)^{-1} > 0$, $\frac{\partial}{\partial K} \overline{\overline{\overline{w_T}}}((K-a)Z)^{-1} > 0$ if $K^2(h(k-a) + \zeta(hk-a)) + \zeta a^2(2K-a)(1-h) > 0$, an opaque expression, negative otherwise. What can be said for sure is that a rise in K lifts up the $W(\sigma)$ schedule and, whenever $a < k$, the $\underline{w_T}(\sigma)$ function, reducing the space occupied by blind compliance regimes in the cases in which $a > k$ and the space occupied by revelation regimes as well whenever $a < k$.

k : *soft target type's marginal cost of compliance*

$\frac{\partial \overline{w_T}(\sigma)}{\partial k} < 0$ if $a < \frac{\sigma}{\zeta} K$, positive otherwise; $\frac{\partial \overline{\overline{w_T}}}{\partial k} < 0$, $\frac{\partial \overline{\overline{\overline{w_T}}}(\sigma)}{\partial k} < 0$, $\frac{\partial \underline{w_T}(\sigma)}{\partial k} < 0$ if $a < \frac{K(\sigma - h(1+\sigma+\zeta))}{\zeta(1-h)}$; this last expression is positive if $h < \frac{\sigma}{1+\sigma+\zeta}$, negative otherwise; $\frac{\partial \widetilde{\sigma}}{\partial k} = 0$, $\frac{\partial}{\partial k} \overline{w_T}((K-a)Z)^{-1} < 0$, $\frac{\partial}{\partial k} \overline{w_T}(kZ)^{-1} < 0$, $\frac{\partial}{\partial k} \overline{\overline{\overline{w_T}}}((K-a)Z)^{-1} < 0$. For low values of a and h, a rise in k causes a drop in all schedules.

a : *sanctioner's marginal gain*

$\frac{\partial \overline{w_T}(\sigma)}{\partial a} < 0$, $\frac{\partial \overline{\overline{w_T}}}{\partial a} < 0$, $\frac{\partial \overline{\overline{\overline{w_T}}}(\sigma)}{\partial a} < 0$ if $h < \frac{K(1+\sigma)}{K(1+\sigma)+\zeta(K-k)}$, positive otherwise; $\frac{\partial \underline{w_T}(\sigma)}{\partial a} < 0$ for $h > \frac{\sigma K - \zeta k}{(1+\sigma)K + \zeta(K-k)}$ with $\sigma K - \zeta k > 0$, positive otherwise; $\frac{\partial \widetilde{\sigma}}{\partial a} > 0$, $\frac{\partial}{\partial a} \overline{w_T}((K-a)Z)^{-1} > 0$, $\frac{\partial}{\partial a} \overline{w_T}(kZ)^{-1} < 0$, $\frac{\partial}{\partial a} \overline{\overline{\overline{w_T}}}((K-a)Z)^{-1} < 0$. If h has an intermediate value, $\frac{\sigma K - \zeta k}{(1+\sigma)K + \zeta(K-k)} < h < \frac{K(1+\sigma)}{K(1+\sigma)+\zeta(K-k)}$, a rise in a brings all schedules down. If h has a high value, $h > \frac{K(1+\sigma)}{K(1+\sigma)+\zeta(K-k)}$, a rise in a expands the range of the $\overline{\overline{\overline{w_T}}}(\sigma)$ function left and right and lifts it up; in the case where $k > a$, the $\underline{w_T}(\sigma)$ function comes down. If h has a low value, $h < \frac{\sigma K-\zeta k}{(1+\sigma)K+\zeta(K-k)}$, a rise in a brings the $W(\sigma)$ schedule down while the $\underline{w_T}(\sigma)$ function, in the case where $k > a$, goes up.

h : *frequency of soft types*

$\frac{\partial \overline{w_T}(\sigma)}{\partial h} > 0$, $\frac{\partial \overline{\overline{w_T}}}{\partial h} > 0$ if $a > k$, negative otherwise; $\frac{\partial \overline{\overline{\overline{w_T}}}(\sigma)}{\partial h} > 0$, $\frac{\partial \underline{w_T}(\sigma)}{\partial h} < 0$ if $a < \min\left(\frac{\sigma}{\zeta}K, k\right)$, which corresponds to the domain of the $\underline{w_T}(\sigma)$ function since $a < \frac{\tilde{\sigma}K}{\zeta}$; $\frac{\partial\tilde{\sigma}}{\partial h} > 0$, $\frac{\partial}{\partial h}\overline{w_T}((K-a)Z)^{-1} > 0$, $\frac{\partial}{\partial h}\overline{w_T}(kZ)^{-1} > 0$, $\frac{\partial}{\partial h}\overline{\overline{\overline{w_T}}}((K-a)Z)^{-1} > 0$. If $a > k$, a rise in h raises up and to the right the entire $W(\sigma)$ schedule. If $a < k$, the same happens except for the $\overline{\overline{w_T}}$ function, which drops and shifts to the right, and the now relevant $\underline{w_T}(\sigma)$ function, which also drops. In both cases, the range of parametric values for which the revelation equilibrium obtains expands.

$-w_T$: *target's cost of war*

Define $\overline{\sigma}(w_T) = \overline{w_T}(\sigma)^{-1}$, $\overline{\overline{\overline{\sigma}}}(w_T) = \overline{\overline{\overline{w_T}}}(\sigma)^{-1}$, $\underline{\sigma}(w_T) = \underline{w_T}(\sigma)^{-1}$; $\frac{\partial\overline{\sigma}(w_T)}{\partial w_T} < 0$, $\frac{\partial\overline{\overline{\overline{\sigma}}}(w_T)}{\partial w_T} < 0$, $\frac{\partial\underline{\sigma}(w_T)}{\partial w_T} > 0$ for the relevant range $k > a$. A rise in w_T, that is a drop in war payoff $-w_T$, makes blind compliance better than revelation or blind defiance, and revelation better than blind defiance.

References

Baldwin D (1985) Economic Statecraft. Princeton University Press, Princeton

Crawford NC (1999) How arms embargoes work. In: Crawford NC, Klotz A (eds) How sanctions work: lessons from South Africa. St. Martin's Press, New York, pp 45–74

Dashti-Gibson J, Davis P, Radcliff B (1997) On the determinants of the success of economic sanctions: an empirical analysis. Am J Political Sci 41(2):608–618

Davidson J, Shambaugh G (2000) Who's Afraid of Economic Incentives? In: Chan S, Drury AC (eds) Sanctions as economic statecraft: theory and practice. St. Martin's Press, New York, pp 37–64

Doxey M (1996) International sanctions in contemporary perspective, 2nd edn. St. Martin's Press, New York

Drezner DW (1999) The sanctions paradox: economic statecraft and international relations. Cambridge University Press, Cambridge

Drezner DW (2003) The hidden hand of economic coercion. Int Org 57(3):643–659

Drury AC (1998) Revisiting economic sanctions reconsidered. J Peace Res 35(4):497–509

Hoffmann F (1967) The functions of economic sanctions: a comparative analysis. J Peace Res 4(2):140–160

Hufbauer G, Schott J, Elliott KA (1990) Economic sanctions reconsidered, vol 2, 2nd edn. Institute for International Economics, Washington DC

Juster KI (2000) The United States and Iraq: Perils of Engagement. In: Haass RN, O'Sulivan ML (eds) Honey and vinegar: incentives, sanctions, and foreign policy. The Brookings Institution Press, Washington, DC

Kaempfer WH, Lowenberg AD (1988) The theory of international economic sanctions: a public choice perspective. Am Econ Rev 78(4):786–793

Kreps DM (1990) A course in microeconomic theory. Princeton University Press, Princeton

Kreps D, Wilson R (1981) Reputation and imperfect information. J Econ Theory 27:253–279

Kydd A (2001) Trust building, trust breaking: the Dilemma of NATO enlargement. Int Org 55(4):801–828
Lacy D, Niou EMS (2004) A theory of issue linkage and economic sanctions: the roles of information, preferences, and threats. J Polit 66(1):25–42
Laffont J-J, Martimort D (2002) The theory of incentives: the sanctioner-target model. Princeton University Press, Princeton
Lopez GA, Cortright D (2004) Containing Iraq: sanctions worked. Foreign Aff 83(4):90–103
Miers AC, Morgan TC (1999) When threats succeed. American Political Science Association, Atlanta, GA (September)
Morgan TC, Schwebach VL (1997) Fools suffer gladly: the use of economic sanctions in international crises. Int Stud Q 41:27–50
National Association of Manufacturers (NAM) (1997) A Catalog of New US Unilateral Economic Sanctions for Foreign Policy Purposes 1993–1996. http://www.usaengage.org/archives/studies/nam2.html
Nooruddin I (2002) Modeling selection bias in studies of sanctions efficacy. Int Interact 28:59–75
Nossal KR (1989) International sanctions as international punishment. Int Org 43(2):301–322
Pape RA Jr (1997) Why economic sanctions do not work. Int Secur 22(2):90–136
Russett BM (1969) Refining deterrence theory: The Japanese Attack on Pearl Harbor. In: Pruitt DG, Snyder RC (eds), pp 127–135
Schwebach VL (2000) Sanctions as signals: a line in the sand or a lack of resolve? In: Chan S, Drury AC (eds) Sanctions as economic statecraft. St Martin's Press, New York
Smith A (1996) The success and use of economic sanctions. Int Interact 21(3):229–245
Verdier D (2008) Successful and failed screening mechanisms in the two gulf wars. J Theor Polit (in press)
Wagner RH (1988) Economic interdependence, bargaining power, and political influence. Int Org 42(3)(Summer):461–484
Yager J (1985) Nuclear supplies and the policies of South Korea and Taiwan toward nuclear weapons. In: Jones R et al (eds) The nuclear suppliers and nonproliferation. Lexington Books, Lexington, pp 187–195