A Theory of Complementary Institutions

Jenna Bednar[1]
Department of Political Science
University of Michigan
jbednar@umich.edu

October 2004

Abstract: To solve problems of compliance, an institutional trigger mechanism is often used. This paper develops a model of two trigger mechanisms and derives conditions where the two institutions may complement one another to improve utility and compliance. When the conditions for optimality do not hold, small changes in institutional structure may create lumpy changes in behavior. Also, the paper explores the consequences of politically-determined institutions. The paper illustrates the model with applications to administrative theory, common pool resource problems, and federalism.

1 Introduction

Guaranteeing compliance in the face of opportunistic incentives is a problem that cuts across the subfields in political science. One solution is to construct institutions that through their monitoring and sanctioning capacities provide counterincentives, boosting compliant behavior. However, in many cases the sanctions necessary to obtain complete compliance are too costly (McNollgast 1987) or may not even be possible, generally (Bednar 2004). Distinct literatures suggest a common solution: multiple sanctioning devices may improve compliance above what may be achieved, feasibly, by a single device. Furthermore, two literatures suggest that by adding an institution capable of only mild sanctions—such as a court—overall compliance may increase compared to when only a severe sanctioning mechanism exists.

Multiple sanctioning institutions play important roles in understanding behavior in political phenomena ranging from control of the bureaucracy to watershed management to federalism. In theories of bureaucratic control, lobbyists and the general public can supplement Congressional oversight of agency behavior (McCubbins & Schwartz 1984, McNollgast 1987, etc). In management of common pool resources, where each participant has an incentive to overuse the resource, Ostrom (1990) and colleagues cite the efficacy of graduated sanctions, with mild punishment for those who shirk minimally, but more severe punishment for those who violate compliance agreements frequently or extensively. Finally, in federalism, no single institutional remedy is sufficient to constrain opportunistic behavior—encroachment, burden-shifting, and shirking—by the member governments. Instead, combinations of institutions interact to counter opportunistic tendencies. These applications will be discussed further below.

The analyses of multiple institutional remedies are suggestive of its efficacy, but have not yet been modeled. With a model we can ask: under what conditions, if any, will multiple institutions increase compliance? When will they improve utility? What are the conditions that lead institutions to fail in their complementarity? In particular, can we *derive* an

effective judiciary?

The potential research agenda is large. In this paper, I establish the efficacy of complementary institutions. I begin by building a baseline model of a trigger mechanism. I then add a piece of a second institution: initially, in Section 3, just a second signal is added. In Section 4 I introduce a full second trigger mechanism and establish conditions when adding a second trigger improves utility and compliance, and use these results to consider the effect of failure to meet these conditions, which may have political roots. Section 5 provides some brief illustrations. Section 6 concludes with a discussion of the meaning of the law.

2   Baseline Model of a Trigger Mechanism

We consider a general game where player utility is a function of the actions of other players, its own shirking, and any sanctions. Although the model will describe a public good provision game, where all players contribute to a common good, it may be adapted to a principal-agent context, where a principal relies upon an agent's dispatch of its responsibilities. We consider games of imperfect information: a player's action is not directly visible to other players or to the sanctioning institutions. Instead, all see a signal of the player's action. In specific games signals may be tied directly to a single player's action, or general, as an indicator of aggregate action. For example, in the Green & Porter (1984) context, the price of oil is an aggregate signal of the oil production of each member of the cartel, while in a Gilligan & Krehbiel (1987, 1989) context, the policy realization or outcome is a signal of the intent of the committee that authorized the policy, and would be indexed to indicate the policy source. In the model described below, each player transmits a signal.

In these games, there is a target effort level from the contributing players: all, in the public good provision game, or any agent, in a principal-agent context. The target effort level is designed to maximize production, and therefore maximize utility for the principal or social welfare in the public good provision game. In either context, we assume that there is a cost for the player to comply, and therefore he prefers to shirk.

A trigger mechanism may induce compliance. In general, the trigger mechanism has

three components: a threshold, which it compares the signal against, a punishment, and a duration of the punishment. In this model, I will collapse the last two elements into a single parameter, a fine, so that the punishment is realized as a utility loss.

Suppose that there are N players and that each must decide how much to contribute to the public good. We will define shirking as a continuous variable $x, x \in [0, 1]$, so that x is an expression of the degree of noncompliance. Let the public good contribution equal $1 - x_i$. I am going to make this model as simple as possible to highlight the effect of complementary institutions. To that end, I am going to assume targeted monitoring, targeted punishment, and fines rather than punishment periods. I am also going to assume that it doesn't cost the players anything to punish those who deviate. This makes for cleaner, more transparent mathematics.

I first consider the case with only one monitoring and sanctioning institution that sets a threshold of T. If the level of shirking equals x_i, then the signal equals $x_i^2 + \epsilon$ where ϵ is uniformly distributed in the interval $[0, 1]$. The squared term implies that small amounts of shirking do not greatly affect the probability that the player is sanctioned. Let Q be the amount of the fine if the signal $x_i^2 + \epsilon > T$. As is standard let δ denote the discount factor. Assume that the single period utility to player i equals

$$u_i(x_1, x_2, ..x_N) = \alpha \sum_{j=1}^{N}(1 - x_j) + x_i.$$

Provided that α is less than one, it is in each player's interest to shirk. Player i then chooses x_i to maximize his value function V which can be written as

$$V(x_i) = K + (1 - \alpha)\, x_i - max\{1 - (T - x_i^2), 1\}Q + \delta V$$

where $K = \alpha + \alpha \sum_{j \neq i}(1 - x_j)$. Solving for V gives

$$V(x_i) = \frac{K + (1 - \alpha)\, x_i - max\{1 - (T - x_i^2), 1\}Q}{1 - \delta}$$

Assume first that $x_i^2 < T$. Then the first order necessary condition equals

$$(1 - \alpha) - 2x_i Q = 0$$

This implies $x_i^* = \frac{(1-\alpha)}{2Q}$. Note that as Q increases x_i gets smaller – the larger the punishment, the less the players deviate. The value to the player from choosing this value of x_i equals

$$V(x_i) = \frac{K + (1-\alpha)\frac{(1-\alpha)}{2Q} - (1 - (T - (\frac{(1-\alpha)}{2Q})^2))Q}{1-\delta}$$

Note that the player is punished if $x^2 + \epsilon > T$, so the probability that the player gets punished equals $1 - T + (\frac{(1-\alpha)}{2Q})^2$.

Alternatively, the player could choose $x_i = 1$ and accept the punishment and get

$$V(x_i) = \frac{K + (1-\alpha) - Q}{1-\delta}$$

So the ability to sustain some cooperation depends upon whether

$$K + (1-\alpha)\frac{(1-\alpha)}{2Q} - (1 - (T - (\frac{(1-\alpha)}{2Q})^2))Q \ <> \ K + (1-\alpha) - Q$$

which reduces to

$$(T - (\frac{(1-\alpha)}{2Q})^2)Q \ <> \ (1-\alpha)(1 - \frac{(1-\alpha)}{2Q}).$$

It is easy to see that a player becomes less likely to shirk fully as T, the threshold that triggers the sanctions, increases. Increasing Q, the penalty levied, has the same effect.

Figure 1 illustrates the logic. The probability of triggering a sanction takes an s-shape. For small amounts of deviation, there is very low probability of crossing the threshold and triggering a sanction, and the curve is relatively flat. For high levels of noncompliance, the probability of crossing the threshold reaches 1. For intermediate levels of compliance, the curve is steeper. Expected utility, graphed in the same figure, initially rises in shirking, peaks, and then begins to fall as the probability of triggering the punishment rises. At some point the expected utility begins to rise again: the player is likely to trigger the sanctions, and so utility increases in degree of shirking, as the player gains utility from its own shirking. Note the trade-off in increasing the frequency of triggering the punishment (or increasing the punishment disutility) and the ability to maintain a cooperative equilibrium:

as punishment becomes increasingly likely, the expected utility from any moderate degree of shirking declines, raising the comparative appeal of the fully noncooperative strategy.

Figure 1 about here

With this baseline model established, we can consider the effect of adding institutions.

3 A Model with Multiple Signals: "And" & "Or"

Our first extension of the baseline model is to consider the effect of a second signal. A second signal may provide insurance in one of two ways: it may be a back-up measure ("or") or it may be a source of verification ("and"). We consider the latter function first.

Consider an additional institution that does nothing more than provide a second look—an additional measure on the signal. Now, if $x_i^2 + \epsilon > T$, rather than immediately triggering the a punishing sanction, let a second institution get its own independently-drawn signal. Let the two signals be s_1 and s_2. Note that since the second signal is drawn from the same distribution (that is, is a result of the same action), its mean & variance are identical, so the signals are correlated. Remember that in equilibrium no one deviates from the common behavior (usually some degree of partial compliance), therefore all signals less than the threshold are due to stochastic error. In equilibrium, it was unlikely before that $s > T$; now, we need both s_1 AND s_2 to exceed the threshold to trigger a punishment. Adding an institution in this way, by requiring a "second look," means that fewer punishments will be triggered. The punishment regime is best thought of as a compliance-maintenance regime; that is, as long as the incentives are adequate, in equilibrium no one deviates, and the punishment is triggered by random environmental noise. When we reduce the frequency of the punishment, the equilibrium behavior induced may be higher, since each player gains more from being in a cooperative state. Therefore a second, identical, complementary institution works on the first two goals, maximizing cooperation and minimizing the cost of enforcement. This is the most common application of redundance: reduction of the Type I error.

On the other hand, if the concern instead is about the likelihood of a Type II error, that the institution might not trigger frequently enough, one might introduce a second signal and trigger a punishment if s_1 OR s_2 exceed the threshold. The frequency of punishment increases. This might be a good solution in the cases where a more severe punishment is not available or the threshold cannot be moved to a more efficient position. (I will discuss further "real-world" limitations on institutional design in Section 5.)

The court, with its hearing of facts, is one institutional means to get a second signal. But it is not necessary; all that is needed, to get a gain in efficiency, is a second, independently-drawn signal. Recall the Green & Porter world of oil cartels. For them, the unique signal is the price of oil: if the price drops, oil producers take it as a signal of a possible increase in supply, caused by the defection of one of the cartel members. Green and Porter did not consider the possibility of two signals, but they could have. For example, a second signal could come from input pricing. If independent truckers and tankers suddenly demand a higher price for transporting oil, it could be due to a higher demand for their services, also a signal that could indicate an oversupply. Or, conversely, if the output price—of oil—drops, but the input price remains steady, then cartel members may be more confident that the price hit a demand shock and no mutual punishment is needed.

The theory of the second signal is related to the results in Sah & Stiglitz (1985, 1986), where they compare two organizational forms, the hierarchy and the polyarchy. The redundance of the hierarchy (the "and" type, in our language) builds in a level of conservatism that rejects more good projects than the polyarchy (the "or" type), but the danger of the polyarchy is that it accepts more bad projects. The implication is that our preference for one type of organizational structure over another would depend upon the relative utility of making a Type I versus a Type II error. In Sah & Stiglitz, this calculation is based upon the relative prevalence of good projects to bad projects, as well as the relative loss from bad projects compared to the gain from good projects.

4 A Model with Two Thresholds: "With"

Next I consider the case of two thresholds. There are three components to this section: first, I develop the model and prove the potential for optimal complementarity. In the second subsection, I invert the optimal complementarity results to show graphically how two institutions may fail to complement one another optimally. In the final subsection, I discuss the consequence of politics, where the parameters of the trigger mechanisms are determined politically, not efficiently.

4.1 Optimal Complementarity

Suppose that a second institution can target punishments and gets an independent signal but that it can only levy minor fines. Formally, what we are doing is making the second institution a fully separate trigger mechanism, with its own threshold and own punishment. It follows that if the second institution's punishment is minor, it must be punishing much more frequently to have any effect and so the second threshold is lower than the first. Here, I assume that the signal is $x^2 + \theta$, where θ is a different random variable that is uniform in $[0, 1]$. Set this threshold equal to t and the punishment equal to q. The lower case letters capture the fact that $t < T$ and $q < Q$. The value function now equals

$$V(x_i) = K + (1 - \alpha)\, x_i - max\{1 - (t - x_i^2), 1\}q - max\{1 - (T - x_i^2), 1\}Q + \delta V.$$

I first assume that the state chooses $x_i^2 < t$. Then the first order condition is

$$(1 - \alpha) - 2x_i q - 2x_i Q = 0.$$

This implies $x_i^* = \frac{(1-\alpha)}{2(Q+q)}$. Notice that the equilibrium level of shirking is lower than in the single institution case, where the equilibrium noncompliance was $x_i = \frac{(1-\alpha)}{2Q}$. This result establishes our first claim:

Claim 1: Without any adjustment to the first institution, the addition of a second, milder institution can only increase compliance.

Our first claim established that the addition of a second trigger mechanism may increase compliance: in the second claim, we consider the effect of a second trigger mechanism on utility.

Claim 2: Without any adjustment to the first institution, the addition of a second, milder institution may increase utility.

The value to the player of shirking x equals

$$V(x_i) = \frac{K + (1-\alpha)\frac{(1-\alpha)}{2(Q+q)}(1 - (t - \frac{(1-\alpha)}{2(Q+q)})^2)q - (1 - (T - \frac{(1-\alpha)}{2(Q+q)})^2)Q}{1 - \delta}$$

We can now compare the utility from two trigger institutions to the one threshold case. This is actually really easy. Suppose that $T = 1$ and $t = 0.5$, $Q = 4$, $q = 1$, and that $\alpha = 0.2$, $\delta = 0.9$ and that $N = 11$. This implies that in the one threshold case, $x_i = \frac{0.8}{8} = 0.1$. If players are symmetric, then $K = 9$. So total value equals

$$V(x_i) = \frac{9 + (0.8)(0.9) - (.01)4}{0.1} = 90 + 72 - .4 = 161.6,$$

while in the two threshold case $x_i = 0.08$. So the total value equals

$$V(x_i) = \frac{9.2 + (0.8)(0.92) - (1/2 + (0.08)^2) - (.08^2)4}{0.1} > 92 + 73.6 - 0.6 - 0.4 = 164.6$$

If t is small relative to T, then the value function looks like a roller coaster with the two thresholds: see Figure 2. The first peak is at $(x^t)^2$. After that the value function falls until t, at which point it starts to rise again, since the smaller fine is already being paid. The second peak is at $(x^T)^2$. In this particular example, the equilibrium degree of shirking is $(x^t)^2$.

Figure 2 about here

We can write the value function as a piecewise function.

- $V(x_i) = K + (1-\alpha)\, x_i - (\max 1, 1 - (t - x_i^2))q - (\max 1, 1 - (T - x_i^2))Q + \delta V$ for $x_i^2 < t$

- $V(x_i) = K + (1-\alpha)\, x_i - q - (\max 1, 1 - (T - x_i^2), 1\}Q + \delta V$ for $x_i^2 \in (t, T]$

8

- $V(x_i) = K + (1 - \alpha)\, x_i - q - Q + \delta V$ for $x_i > T$

Using these equations, we can derive the condition necessary for the mild institution to complement the severe:

Complementarity Condition: If $\max V(x_i; 0 < x_i^2 \leq t) > \max V(x_i; x_i^2 > t)$ then the addition of the mild institution improves utility.

Informally, in the illustration, whether the mild institution improves utility depends upon whether the first "hill" is bigger than the second. With two institutions, each with its own trigger mechanism, you have two distinct probability curves. Together, these institutions improve the efficiency of the union: the mild institution sustains compliance at higher levels, and minimizes the frequency of punishment regimes that are triggered, while the continuation of the punishment regime threat—the severe institution—dampens the utility gained from higher levels of cheating, and so prevents large deviations.

The utility and compliance improvements offered by the second, mild institution provides the opportunity to reduce the likelihood of triggering the severe mechanism. The threshold to trigger the original, more severe punishment can be adjusted upward, so we have $Tt < T < TT$, where T is our original threshold, when only one sanctioning institution was available. For example, suppose that the original threshold to trigger the severe punishment was .40. It might be efficiently raised to .60 with the addition of a milder sanction that triggers more frequently, perhaps at perceived deviations of .30. Now, it is the mild institution that triggers and sanctions for any signal between .30 and .60, and the more severe institution that triggers for signals greater than .60, perhaps in conjunction with the mild sanctioning institution.

One may further permutate the combination of institutions to model an adjudicatory mechanism. We use the same illustrative values as above, but now, suppose that the initial signal is perceived by the players, who may react severely to the rare extremely high signal, but for any signal between .30 and .60, they take the perceived deviator to court. The court

then draws its own, independent signal from the same original action, and punishes only if the signal exceeds its threshold.

This game may be further modified to allow for distinct thresholds at the mild end. For example, the players may challenge in court any signal that is greater than .20, but the court only sanction signals it receives that are greater than .30. The players own challenging threshold is crossed more frequently than the court's: the players have lower tolerance for perceived deviance than the courts. Or, one might think of it as the courts having a higher burden of proof to sanction.

We may now consider how one might design the optimal mild trigger, for example, the judiciary. Let's consider the severe institution as fixed and adequate. That is, it is sufficient to sustain some degree of compliance, although not full compliance. For a given T, Q combination, as well as an α, we may consider the optimal placement of the court's trigger threshold and the amount that it fines. This calculation will be highly dependent upon the functional form assumed in the model.

Notice that the institution has bite when its probability of punishment curve is steep. As long as it is steep, the marginal utility from shirking is decreasing. Institutions are efficiently complementary when their prob of punishment curves cover as much of the compliance space as possible. This observation has important implications for the optimal design of institutions. In the next subsection, I illustrate by demonstrating the effect when these conditions are absent.

4.2 Insufficient Complementarity

When the Complementarity Condition is not met, the addition of the mild institution fails to improve utility or compliance. If the severe institution was adjusted with the addition of the mild institution (by raising T or reducing Q), then utility and compliance may *decrease* with the institutional change. In this subsection, we examine two forms of insufficient complementarity: (1) an overly mild institution and (2) an overly weakened severe institution.

The Insufficient Mild Institution

Recall Claim 2, establishing the utility improvement from the addition of the mild institution. Farther along in the discussion, I mentioned that the addition of the mild institution may allow an adjustment to the severe institution, to trigger less frequently, gaining the same amount of compliance with less cost. However, if the mild institution is too mild, a player may choose to ignore it, accepting its penalty. Returning to the example developed above, the player chooses $x_i = \frac{1-\alpha}{2Q}$ (because the institutions were set to keep compliance at the same level) and pays the fine of q. If that is the case, the player would be worse off from the two thresholds.

Figure 3 about here

Figure 3 graphs the outcome from the addition of an overly mild institution. The dotted line represents the probability of triggering the mild institution. In this graph, the Complementary Condition fails to hold, and in fact the value function is maximized in the interval $t < x_i^2 < T$. The mild institution is triggered with certainty, and the player is penalized q, with a potential loss of utility over the single-institution case. Alternatively, the mild institution may trigger too infrequently (t may be set too high) for the threat of punishment to deter the lure of moderate shirking.

In these cases, while a mild institution is present, it has no effect on compliance and can only decrease utility. This example serves as a warning to institutional designers: if a mild institution is likely to be ineffective, it is better not to include it in the institutional plan.

The Insufficient Severe Institution

Figure 4 illustrates the opposite case: while the mild institution has some effect, the severe institution fails completely to motivate the players. Notice the relatively high expected utility in the interval $0 < x_i^2 \leq t$; the mild institution is deterring moderate shirking. However, in the interval $t < x_i^2 \leq T$ the severe institution hardly manages to display any effect. Expected

utility is maximized in the inverval $x_i^2 > T$, specifically, at full non-compliance.

Figure 4 about here

This illustration displays the primal importance of an effective severe institution. If the mild institution, such as a court, were acting alone, it could not sustain an equilibrium of high compliance. Recall the comparative statics at the end of Section 2: a single trigger mechanism is more likely to sustain some compliance (rather than full noncompliance) as T or Q increase. A mild mechanism works in exactly the opposite way: with a low threshold it triggers frequently, but with a low sanction. With such a weak punishment, as a player becomes highly likely to trigger the punishment regime through its noncompliance, it might as well deviate fully. The institution is not able to sustain compliance. The severe sanctions are needed to prevent major deviations. The mild mechanism can stop small cheaters but only the threat of severe sanctions can prevent big deviations.

4.3   Inconsistent Thresholds: Efficiency vs. Politics

We finally consider a case where the thresholds may be inconsistent. That is, to this point we have considered the institutions to be impersonal, disinterested trigger mechanisms, even employing the language of "construction" to make them appear to be distinct from strategic players. However, these institutions are themselves composed of strategic actors, and any sanctioning reaction must be self-enforcing. We continue to assume that the threshold and punishment are exogenously set, but now let us consider a political derivation of these parameter values. For example, perhaps the mild institution is a court, and the severe institution the threat of civil war. The court sets its threshold for reasons external to this model: strategic, behavioral, or legal reasons. But the forces that act upon the court may change, causing its threshold value to shift. If the institutions have inconsistent threshold settings, or generally are set for reasons only indirectly related to maximization of the social welfare function, then the institution may be inefficient, and even ineffective.

Recall, for example, Figure 3. Here the mild institution is not set at a level that is

effective, and it is ignored. It is easy to see that a small shift in the mild institution may cause a lumpy shift in behavior, as equilibrium levels of shirking shift from approaching full compliance, to that sustainable by the severe trigger alone.

One solution to the problem of inconsistent thresholds is the inverse of original motivation for the second signal, in Section 3. There, the second signal served as a confirmation of the first, as a way to minimize the amount of punishment. This is helpful when punishment is costly to those punishing. However, in the present case, the concern is that the signal may fail to trigger a mild sanction. In this case, it is useful to have an "or"-type independent trigger not as confirmation, but as insurance.

Equivalent logic may apply to the severe institution.

Summary of results:

1. Two institutions can improve compliance over a single institution, but not universally. In particular, as long as the original severe institution is not adjusted, it will never decrease compliance.

2. Two institutions can increase utility, but not universally.

3. If the mild institution's threshold is raised to too high a standard, it can have no effect.

4. If the severe institution is adjusted when the mild institution is added, to reduce the frequency of severe punishment, utility may increase. However, the combination of institutions must be adjusted precisely, or the design may backfire, reducing overall compliance levels.

5. Mild institutions should not be studied in isolation: to understand their effect, one must know how they complement more severe institutions.

6. Multiple mild institutions are particularly useful when their thresholds are inconsistent. (Created when the institution has interests of its own, causing it to set T and Q at inefficient levels, from the social welfare viewpoint.)

5 Applications

DEAR READER, THIS SECTION IS UNDER DEVELOPMENT.

5.1 Administrative Law: Police Patrol and Fire Alarms

> Because monitoring and enforcement are not costless, no method of influencing
> administrative decisions will be perfect. Rather, elected representatives face a
> tradeoff between the extent of compliance they can command and the effort that
> is expended to assure it, effort which has an opportunity cost because it can also
> be used for other politically relevant purposes (McCubbins, Noll, & Weingast
> 1987:247).

Congress delegates authority because it cannot monitor every act an agency commits. However, it can do so knowing that there is a back-up monitor: the public and interested organized groups. The public and interest groups can watch the activity of the agency and sound an alarm, either by vocalizing complaints to Congress, or, less costly to Congress, by challenging the activity in court. (congress thereby exports the costs of enforcement.) The literature in this field is well developed (for a recent review, see the Bendor, Glaser, & Hammond, 2001?).

There are two advantages of combining police patrols and fire alarms. First, the police patrol, as a metaphor for vigilant monitoring by Congress of agency activity, is costly. By designing administrative law in a way that encourages the public to monitor the activities of an agency, by giving it a legal remedy for perceived violations, the cost of monitoring is reduced. By itself this would not necessarily increase compliance. However, it is the second advantage of the fire alarm that I want to emphasize: the existence of a second monitoring & sanctioning institution provides a chance for a second signal. It is a straightforward application of the logic developed in section 4.3: if the political establishment of the institutional parameters causes concerns about efficient enforcement, it is best to build a redundant system, where the various institutions fail for distinct reasons. This independence may be

encouraged by having institutions that are influenced by different political processes.

NOTE TO SELF: the argument about administrative law makes conditional statements about when Congress would write laws to facilitate legal challenges only when it benefits Congress to do so.

5.2 Common Pool Resource Problems: Graduated Sanctions

quote from Ostrom.

What is this mild sanction doing? What is its deterrence effect? is it failsafe? Big point: mild not effective if severe didn't exist. also, mild may be miscalculated. Not likely in CPR problems, where there's a close match between the setting of the threshold and the social welfare objective.

5.3 Federalism: Judicial v. Political Safeguards

We value federalism because it creates a public good. To be more specific, its decentralized but unified structure makes it possible for citizens and member states to achieve a wide variety of "meta" public goods, including national security, domestic peace, free trade among members, local sovereignty and policy experimentation, and improvement to democracy. These are rightly considered public goods: the benefits, such as they are, accrue to all members of the community without discrimination, while whatever burdens necessary to provide the public good are shared by all.

At the same time, the union generates private goods to be distributed among its members. These private goods are an aspect of federalism to the extent that they are (1) generated by the federal union, (2) their quantity affects the utility that member governments and the public receive from the union, and (3) their distribution is affected by the nature of federalism specifically, and not just geographical constituencies more generally. For the study of other distributional aspects of private goods, I defer to the fine literature on pork in Congress, as well as the forthcoming book by Filippov, Ordeshook, and Shvetsova, which focuses on the allocation of private goods within a federation.

Instead, here we focus on minimizing opportunism and maximizing the social welfare. The federation is built with various safeguards: Madison emphasized fragmentation of the center, state supervision, including involvement in federal decision-making as well as whistle-blowing, and electoral control. More recently, theorists have added political and judicial safeguards to the list of institutions that might sustain federalism. While federalism is known formally by a spare definition of decentralized government that does not distinguish it from many confederal relationships,[1] it is typically associated so closely with many institutional features, such as bicameralism, equal representation of the states, and often, an independent judiciary, that a complex system of institutions seems inevitable, if not integral, to any successful federation.

This analysis suggest that the institutional framework that supports a federation is not only complex, but its components are interrelated, are *interdependent*. Furthermore, calculation of their efficacy is not as straightforward as in many examples of Common Pool Resource management. Judicial thresholds—what behavior the court will tolerate—and especially political safeguards set their thresholds for reasons not uniquely related to maximization of the social welfare. That is, while the court and political parties may want the federation to be as productive as possible, they have other motivations which might compete with this goal, and therefore the threshold may be set at an ineffecient level. For example, they may be too tolerant of opportunistic behavior—say, encroachment by the federal government—due to political/electoral motivations (in the case of political safeguards) or lack of independence (for judicial safeguards). In either case, the threshold of tolerated behavior may be set well beyond the efficient point; either mild safeguard may simply tolerate more opportunistic behavior than would produce the optimal federal union.

In particular, if the mild institution's ability to sanction is limited, it is critical that the threshold be set to trigger frequently enough that this mild sanction would have an effect. For example, suppose that all that courts can do is declare legislation unconstitutional, nullifying it. Congress can respond by repassing the legislation plus some threat of retribution

---

[1]Riker's definition here.

16

should the court strike it down, or the people can amend the constitution. These are somewhat costly acts, the second more so than the first. But if the court does not question behavior frequently enough, its effect is lost. Notice that this argument contradicts a standard hypothesis: that if the court intervenes too frequently, its effect is diluted, as it loses its legitimacy.

THIS ARGUMENT WILL CONTINUE, MAKING THE FOLLOWING POINTS:

1. The Judiciary as well as political safeguards depend upon a severe back up to be effective. In particular, we may blame the court for being too weak to be effective, when it is not designed to stop significant opportunism. When a credible threat of severe sanctions exists (ie civil war!), then the court may be able to improve compliance—and perhaps even increase utility—by deterring minor transgressions. Mild sanctions can do virtually nothing on their own, without a complementary severe trigger mechanism. (I AM THINKING THAT THIS NECESSITY MIGHT HELP US TO UNDERSTAND THE IMPORTANCE OF THE CREATION OF A 'NATION' IN SUSTAINING FEDERALISM—ESTABLISHMENT OF THE NOTION THAT THE UNION IS INDIVISIBLE, OR AT LEAST NOT WITHOUT A GOOD FIGHT. Nationhood, or a federal culture, may make a severe threat credible.)

2. Judicial and Political safeguards complement one another: they each have inconsistent thresholds, as they are set for reasons external to the model, and so if you want to be sure that at least one mild institution is triggered, it is best to have redundancy. (Notice how this works contrary to the standard redundancy argument, where you want a "second opinion," confirming the diagnosis before triggering. Here, the concern is that you wouldn't trigger often enough, and so you want redundant triggers.

3. However, the advantages of redundancy are qualified: recall that an ineffective mild institution may reduce overall utility. If the structure of institutions in a federation are on net utility-decreasing, the exit option of union dissolution becomes comparatively more attractive. This caution may bolster arguments to limit the intervention of certain

institutions, such as the court, in federalism disputes. Any argument in support of the court will have to be able to claim that its involvement as a complement to political safeguards enhances utility, rather than decreases it; the force of such an argument will depend upon a compelling case made for the independence of the errors between the judiciary and the political process.

4. We can begin to look at realms of law where one safeguard may be more effective than another: for example, distinguishing between patrolling spending powers, which the court is terrible at, may be excusable because of less tolerant political safeguards. On the other hand, the court has become increasingly interventionist in commerce clause issues, where political safeguards are pathetically tolerant of federal encroachment. *Lopez* and *Morrison* are good examples here.

6   Discussion

This analysis suggests that to focus on the remedial potential of a single institution is incomplete in many contexts: it is partial equilibrium analysis. In order to understand how an institution works, its potential to induce compliance, and especially its design, we must consider the full institutional context. We see that mild sanctioning devices, such as a court, may be useful at dissuading minor acts of noncompliance, but it is unable to prevent major acts—it is not designed to do so! Its efficacy, even at preventing minor shirking, depends on reinforcement by a more severe sanctioning force. The compliance increase and any utility increase is not only a function of the positioning of its threshold and severity of its punishment, but also, these qualities in relation to the more severe trigger mechanism.

This research invites multiple types of extensions. One potential is to consider the complementarity in institutions when the behavior to be modified is more complex. For example, in my own substantive field, federalism, institutions may prevent opportunism, but opportunism comes in three forms, typed by perpetrator and victim: the federal government may *encroach* upon the states, the states may *burden-shift* on one another, and the states may

shirk on their duties to the federal union. We would not necessarily assume that the same institution would constrain all three forms of opportunism. Instead, institutions may complement one another in coverage of the three types. This example introduces a third form of institutional complementarity to be considered in future research.

References

Bednar, Jenna. 2004. "Is Full Compliance Possible?" Manuscript under review; manuscript available at www.umich.edu/ jbednar.

Bendor, Glaser, & Hammond.

Eskridge, William N., Jr. & Philip P. Frickey. 1994. "The Supreme Court, 1993 Term Foreword: Law as Equilibrium." *Harvard Law Review* 108:26-ENDPAGE.

Gilligan, Tom and Keith Krehbiel. 1987. "Collective Decision-Making and Standing Committees: An Informational Rationale for Restrictive Amendment Procedures." *Journal of Law Economics and Organization* 3:287-335.

Gilligan, Tom and Keith Krehbiel. 1989. "Asymmetric Information and Legislative Rules with a Heterogeneous Committee." *American Journal of Political Science* 33:459-490.

Green, Edward J. and Robert H. Porter. 1984. "Noncooperative Collusion under Imperfect Information." *Econometrica* 52(1): 87-100.

McCubbins, Mathew, Roger Noll and Barry R. Weingast. 1987. JLEO.

McCubbins, Mathew D. and Thomas Schwartz. 1984. "Congressional Oversight Overlooked: Police Patrol versus Fire Alarms." *American Journal of Political Science* 2(1):165-179.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action.* Cambridge, UK: Cambridge University Press.

Sah, Raaj Kumar and Joseph E. Stiglitz. 1985. "Human Fallibility and Economic Organization." *American Economic Review, Papers and Proceedings* 75(2):292-297.
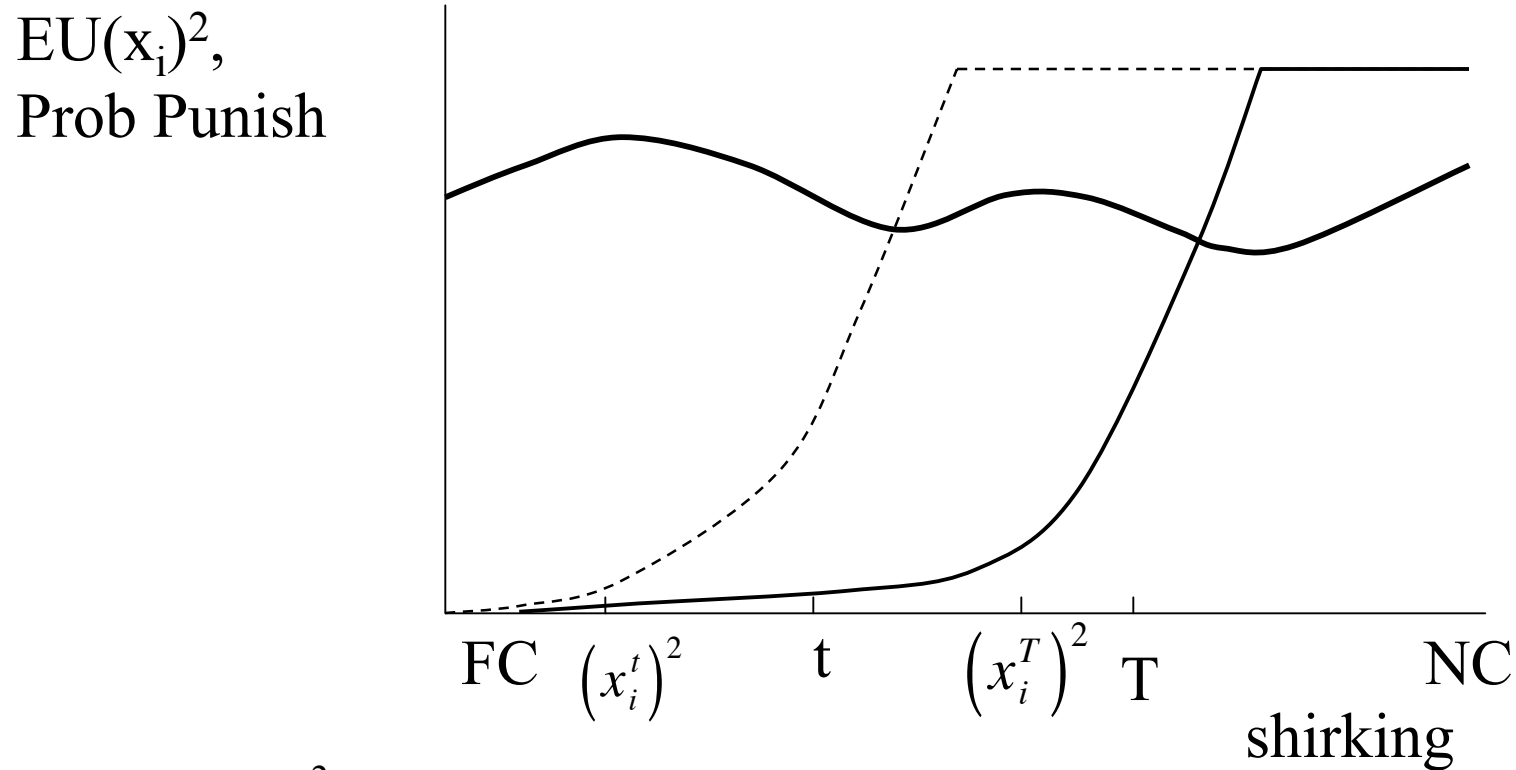
Sah, Raaj Kumar and Joseph E. Stiglitz. 1986. "The Architecture of Economic Systems: Hierarchies and Polyarchies." *American Economic Review* 76(4):716-727.

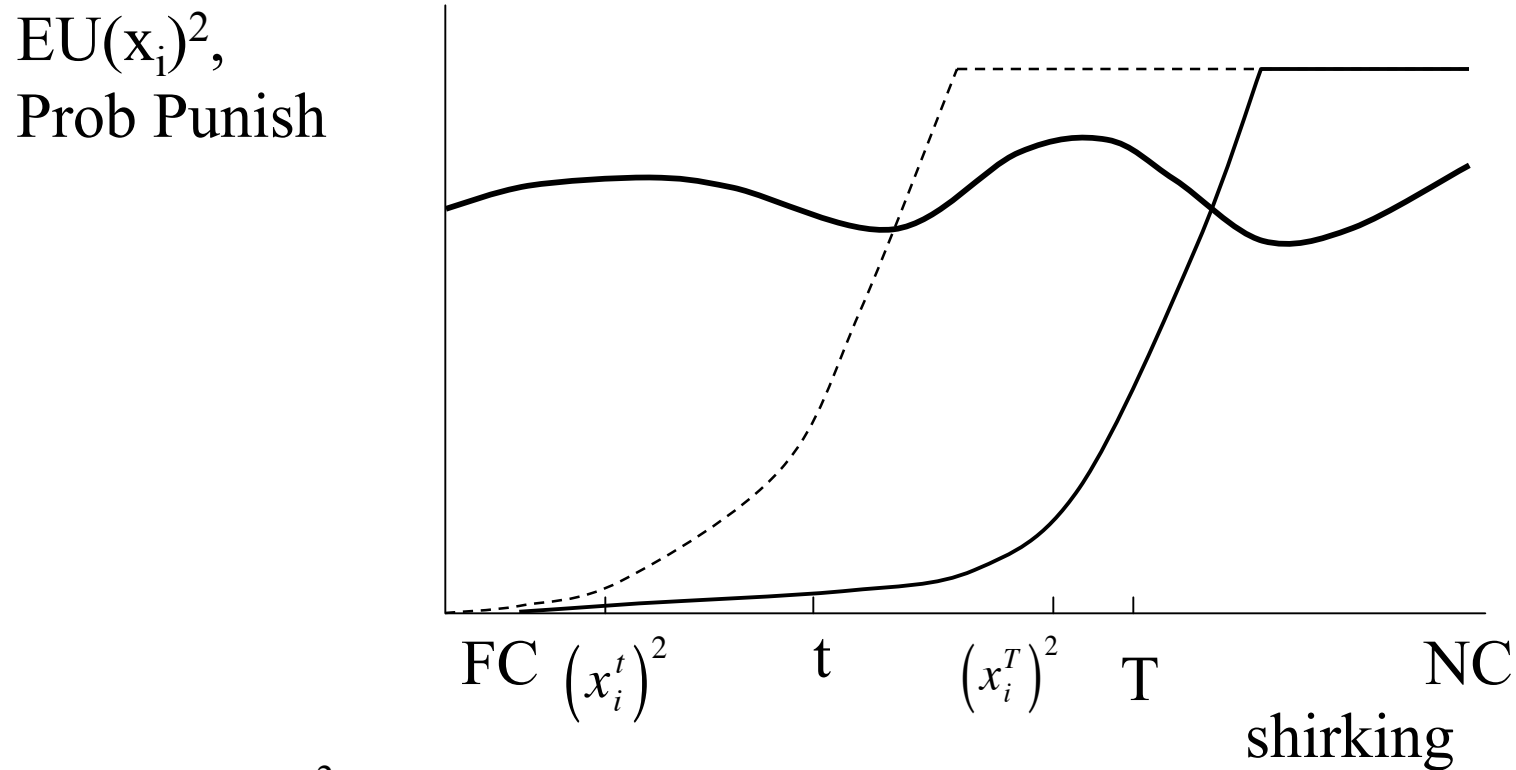Figure 1: Baseline Model; Inherent Opportunism
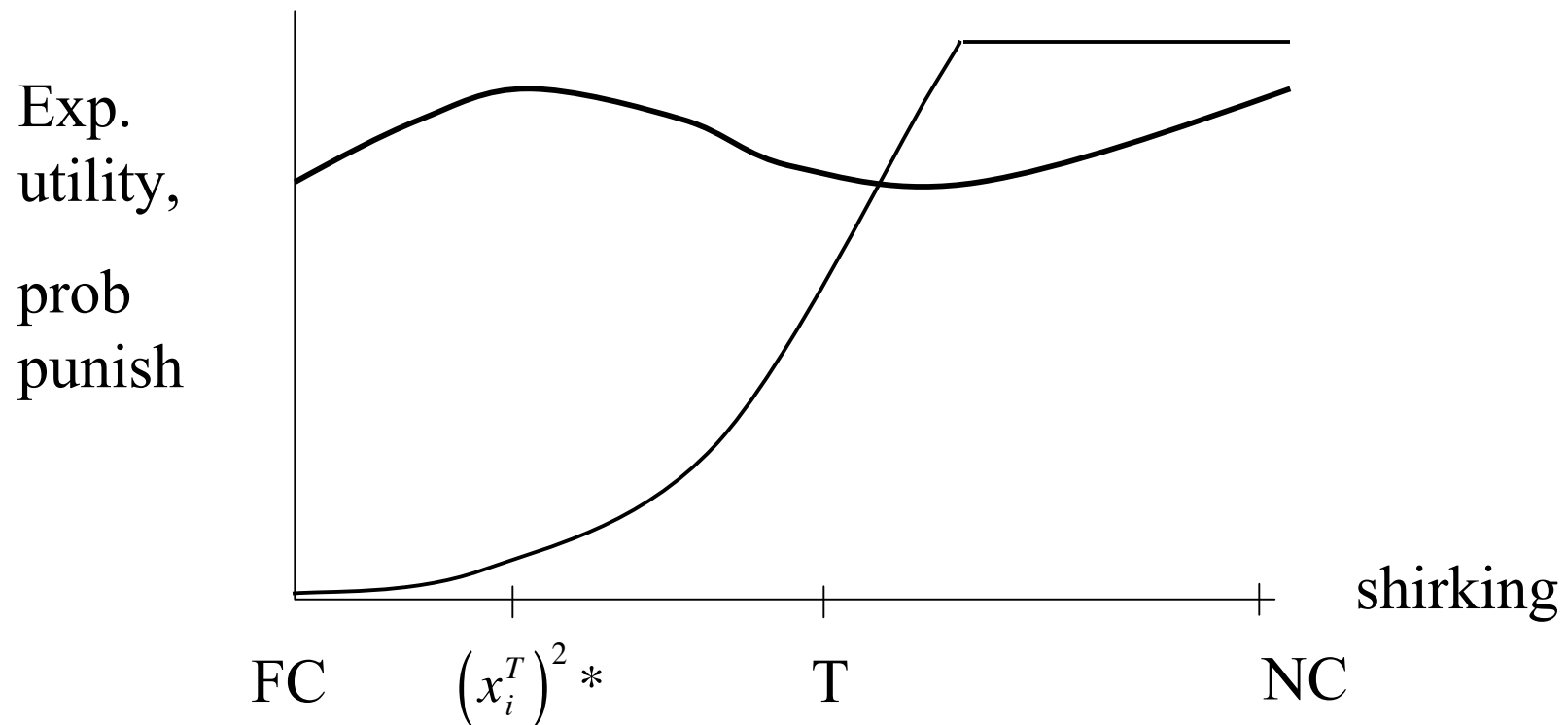
$\mathrm{EU}\left(x_i^T\right)^2 * > \mathrm{EU}(\mathrm{NC}):$ partial compliance equilibrium

Figure 2: Complementary Institutions

EU(x_i)^2,
Prob Punish

FC $\left(x_i^t\right)^2$ t $\left(x_i^T\right)^2$ T NC

shirking

Where $\left(x_i^t\right)^2 = x_i^2 *$

Figure 3: Insufficient Mild Institution

$EU(x_i)^2$, Prob Punish

FC $\left(x_i^t\right)^2$ t $\left(x_i^T\right)^2$ T   NC

shirking

Where $\left(x_i^T\right)^2 = x_i^2 *$

Figure 4: Insufficient Severe Institution

$EU(x_i)^2$,
Prob Punish

FC   t   T x^2*=NC
                                   shirking