COLLECTIVE SANCTIONS AND COMPLIANCE NORMS: A FORMAL THEORY OF GROUP-MEDIATED SOCIAL CONTROL

DOUGLAS D. HECKATHORN
University of Missouri-Kansas City

The link between external sanctions and intragroup normative control is examined to distinguish the conditions under which the two control systems augment or weaken one another. I construct a dynamic rational choice model that incorporates essential features of the sanction/norm link. The analysis suggests that much social control that appears to result from the sanctions of individuals derives instead from a form of group-mediated control termed compliance norms. In other contexts, intragroup control acts in opposition to external sanctions, resulting in a form of control termed opposition norms. According to the formal analysis, group responses to sanctions depend on the strength of sanctions, monitoring capacities, and the efficacy and cost of intragroup control.

Efforts to understand behavioral responses to sanctions have given rise to an extensive literature in the fields of criminology and deviance (Andenaes 1974; Meier and Johnson 1977; Pilavin, Gartner, Thornton, and Matsueda 1986), social psychology (Felson and Rusco 1988), and economics and political science (Becker 1976). This literature demonstrates the complexity of the underlying processes. Studies of the deterrence doctrine are a case in point. As these studies became methodologically and theoretically more sophisticated, the limitations of the classical individualistic deterrence model became increasingly evident (Gibbs 1975; Brier and Fienberg 1980). For example, contrary to the deterrence doctrine, criminal organizations may compensate for lost income by increasing their level of criminal activity. They may respond to increased apprehension of their members by accelerating the recruitment of replacements. If a criminal organization responds to enforcement efforts by seeking to preserve its operations, the effect may be a pattern of response to sanctions that is the opposite of that of a rational individual faced with comparable sanctions. Thus, studies of responses to sanctions must take into account the group context within which sanctioning occurs (Gibbs 1975; Heckathorn 1985).

Most studies continue to focus on strictly individualized sanctions, examining rewards for compliance or punishments directed at violators. Only a few analysts (Heckathorn 1988) focus on collective sanctions — systems where rewards or punishments extend not only to the actor but to the actor’s group. In such systems, groups’ responses to the external sanctions are highly variable. In some cases, the group complies by creating a secondary sanction system — termed exogenous compliance norms (Heckathorn 1988) — that reinforces the original sanction. The result can be extraordinarily high levels of compliance because the groups’ internal monitoring capacities and sanctioning resources are harnessed on behalf of external control. In other cases, the group resists efforts to control the group. As a result, the group may act against the external agent that controls sanctions. Thus, collective sanctions create ambivalent incentives (Heckathorn 1988, p. 556). But, whether the group reacts compliantly or rebelliously, their responses to collective sanctions are themselves collective.

Social sanctions can be arranged along a continuum from strictly individual sanctions that target a single individual to strictly collective sanctions that impact equally upon both the actor whose behavior triggered the sanction and the group of which that actor is a member.

*Earlier versions of this paper were presented at the Theory Workshop of the University of Iowa, the Public Choice Society meetings, and a Liberty Fund conference. The paper benefited from comments by many participants. Especially thank Susan LoBello, Debra Friedman, Michael Hechter, Pamela Oliver, Hartmut Kliemt, Roger Carlson, and the reviewers of ASR for their valuable comments and advice. I would also like to acknowledge support provided by the George A. and Eliza Gardner Howard Foundation.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
At intermediate points on the continuum lie what may be termed mixed sanctions. These are individual sanctions which possess externalities that spill over to the actor’s group.

In the real world of social behavior, few sanctions are either strictly individual or strictly collective. Virtually all individuals are members of groups such as, family members, friends, neighbors, co-workers and others with whom the individual interacts. To the extent that members of a group are interdependent, sanctions directed at any individual have consequences for other group members. For example, when an individual is arrested or fired, that event also sanctions family members and friends. Except in the limiting case of social isolates, virtually all sanctions generate externalities.

Social sanctions are rarely strictly individualized. In some cases, spill-over effects are limited and can be ignored by analysts. Such is the case for very weak microsocial sanctions. For example, when an individual is criticized by peers, the sanction’s effects are localized and transient. In contrast, a stronger microsocial sanction, such as expelling miscreants from a group, has a significant systemic impact since it reduces the group’s size.

The spill-over from individual to collective sanctions derives from three fundamental sources. First, it derives from the interdependence of group members. In groups that are “positively connected” (Cook and Emerson 1978), a reward received by one individual benefits other group members, so the individual sanction gives rise to a secondary collective sanction. This occurs either because of material interdependence within the group or because of altruistic orientations of group members. Many social groups such as familial and work groups are in some measure positively connected, so this source of dispersion effects is significant.

Second, individual sanctions generate externalities via equal treatment norms or expectations. For example, if an employer is expected to treat all workers uniformly, a unilateral concession by one worker can pressure others to make the same concession. That provides the basis for “output restriction” norms (Mathewson 1931) that forbid “rate-busting” and other forms of over-compliance to authority. The expectation of equal treatment can also generate collective sanctions outside of organizations. For example, a rash of crimes in a neighborhood frequently makes its residents fear that they may be the next target (Silberman 1978). Each crime punishes not only the victim, but other neighborhood residents as well.

Finally, secondary collective sanctions derive from an explicit system of collective reward or punishment wherein group members are held accountable for one another’s conduct. This occurs in the group incentive plans adopted by many progressive U.S. corporations (Miller and Schuster 1987), in military boot camps where one recruit’s violation causes everyone in the barracks to be punished (Gilham 1982), and in tribal societies where concepts of responsibility are generally corporate rather than individual (Driberg 1970).

The mixed character of most social sanctions has important implications for understanding the relationship between macrosocial control processes and microsocial norms. The spill-over from these sanctions gives group members a stake in regulating one another’s behavior. It creates within the group the ambivalent incentives characteristic of collective sanction systems (Hekathorn 1988, p. 556). On the one hand, group members have incentives to urge one another to seek out external sources of rewards and to comply with external dictates to avoid triggering externally induced punishments. In this case, intragroup normative controls amplify the effects of the sanction. On the other hand, peer group members may also have incentives to help violators avoid detection and to assist rebels in their struggle. Here intragroup normative controls weaken and oppose the external sanction.

I examine systems consisting of an agent who provides external sanctions and a group that is endowed with an internal control capacity. I focus primarily on group responses to sanctions, not on the agent’s decisions. Consequently, the nature of the agent is not specified in detail. Whether the agent is the state, a subgovernmental organization, a larger normative system that impinges on the group, or nature, the agent is merely assumed to have a specified set of sanctions and a fixed strategy for deciding when to dispense them.

My substantive goal is to differentiate the conditions under which external sanction-based control and intragroup normative control augment or weaken one another. The approach is to construct a dynamic rational choice model (Coleman 1986b; Hechter 1983; Opp 1988) that incorporates the essential features of the sanction-norm link, and to identify the resulting...
equilibrium patterns of behavior resulting from alternative combinations of system parameters.

COMPLIANCE NORMS AND THE EMERGENCE OF INTRAGROUP CONTROL

When the analysis shifts from individual to group behavior, it is necessary to consider three factors that are irrelevant to strictly individualistic models of control. First, the analysis must consider not merely the sanction targeted at the individual, but also the collective sanctions that, owing to the original sanction’s spill-over effects, target all group members. Second, it must consider the group’s capacity to create and enforce compliance norms that amplify the effects of the individual sanction. Finally, the analysis must consider opposition norms that can arise to oppose that control.

When the spill-over from an individual sanction subjects group members to a collective sanction, the group faces a collective action problem. Collective sanctions create a special form of public goods problem (Heckathorn 1988) owing to the conflict between the collective interest in producing the public good (i.e., interests in collective rewards or avoiding collective punishments), and the individual interest in free-riding (i.e., evading the compliance costs involved in reducing the risk of sanctions). The collective action problem faced by the group can be modeled as an iterated n-person prisoner’s dilemma (Heckathorn 1988, pp. 441-45).

Collective sanctions also create a second collective action problem of creating and enforcing the system of compliance norms. “A sanctioning system is also a public good because its benefits can be enjoyed by all members regardless of their contribution to its provision” (Yamagishi 1986, p.110). Thus arises what Oliver (1980) has termed the second order free rider problem — the potential for actors to enjoy the fruits of the compliance norms (i.e., avoidance of group sanctions) without bearing the costs associated with their creation and maintenance.

The choice for actors in the traditional prisoner’s dilemma is twofold (Rapoport and Chammah 1965): to cooperate or to free ride. However, when collective action includes a second order free rider problem, each actor faces a fourfold choice (Heckathorn 1989, pp. 81-2) since each must choose whether to cooperate at the ‘first level’ of the original collective action problem and at the ‘higher level’ of the intragroup sanctioning system. Full cooperation involves both obeying the agent who controls the sanctions and sanctioning those who disobey the agent. Here the individual complies to prevent triggering sanctions (personal compliance), and exercises control to reduce the risk that others may trigger collective sanctions (second level compliance). The second possibility, hypocritical cooperation, occurs when an actor violates the agent’s dictates while urging others to obey them. That is, the actor defects at the first level but cooperates on the second level. Recent news stories involving television evangelists provide vivid examples of this alternative. The third possibility is private cooperation, cooperating at the first level but defecting at the second. Here the actor obeys the agent, sparing the group the risk that he or she might trigger collective sanctions, but fails to prevent others from violating the agent’s dictates. Finally, the actor can choose full defection, defecting at both the first and second levels — he or she neither personally conforms, nor encourages others to conform.

Basic Assumptions of the Formal Model

In constructing a formal model, I assume that sanctioning is structured as follows: The external agent possesses a monitoring capacity, M, defined as the probability that any single violation of its dictates will be detected. If the agent observes a violation, it administers the individual sanction to the culprit, and the collective sanction to all members of that actor’s group. When sanctions are negative, the strength of the individual sanction, Sj, can be defined as its disutility for the targeted actor. Similarly, the strength of the collective sanction, Sj, can be defined as its disutility for members of the actor’s group. Finally, the extent of spill-over from the individual to the collective sanction can be defined as the ratio of these two sanc-

---

1 The effects of the shape of the production function for the collective good have been intensively investigated (Bonacich, Shure, Kahan, and Meeker, 1976; Oliver, Marwell, and Teixeira 1985). Consequently, sophisticated analyses of collective action now consider the effects of a variety of alternative functions. In studies of collective sanctions, this is unnecessary because the structure of the sanctions uniquely specify the shape of the production function (Heckathorn 1988).
tion strengths, $S_1/S_2$.

Four simplifying assumptions structure this formal model of the sanctioning process:

1. A single detected violation suffices to trigger sanctions. This assumption is not very restrictive, since the effects of greater leniency on the agent's part can be incorporated into the model by attributing a diminished monitoring capacity (M) to the agent. In essence, the model fails to distinguish between cases where the agent detects but ignores violations, and cases where the agent fails to monitor violations effectively. Since I focus on group responses to sanctions rather than on the decision problem faced by the external agent, this distinction is not crucial.

2. Collective sanctions impact equally on all group members. This assumption converts into a step function the common situation in which the strength of the externalities generated by individual sanctions is a sharply diminishing function of social distance from the targeted actor. For example, when a member of a nuclear family is fired, all family members are affected by that event, whereas neighbors are only trivially affected. In an extended family system, these externalities might not be adequately modeled using step functions, and might instead require gradients wherein the strength of the collective sanction is a function of social distance from the actor who triggered sanctions. Space does not permit a consideration of these more complex systems.

3. Both individual and collective sanctions entail the threat of punishment rather than the promise of reward. Hence I consider only negative sanctions. However, as Oliver (1980, p. 1361) states, for rational actors, "positive and negative incentives are radically different in the view of the person who uses them, even though they are the same to the persons receiving them." I focus not on the agents who dispense sanctions (except to acknowledge their existence and specify a single fixed strategy) but only on the behavior of the recipients of sanctions.

4. Groups are composed of homogeneous actors. Consideration of the effects of heterogeneity are beyond the scope of this paper (see Heckathorn 1990a). Specific effects of heterogeneity are noted at several points.

In actual social groups, intragroup control employs a diverse array of mechanisms, including persuasion, wherein actors affect one another's preference orders, selective incentives such as promises and threats that affect the anticipated payoffs from alternative courses of action, and control of opportunity structures wherein opportunities to engage in group approved actions are expanded, and opportunities to engage in disapproved actions are limited. A comprehensive model of intragroup control would encompass these and many other mechanisms of control. However, these complexities need not detain us here since my concern is not with these mechanisms, but with the conditions under which actors are motivated to participate in the creation of systems of intragroup control. Consequently, my model of intragroup control has only three requirements. First, actors must possess a rigorously specifiable capacity to affect another's behavior. Second, exercising such control must entail a specifiable cost to each participating actor. Third, control structures must be treated as social products, with each individual actor's contribution having a specifiable impact on that structure. Hence, the model requires an aggregation procedure by which each individual's exercise of control combines with others into a group control structure. A behavioral control model based on persuasion, incentives, or opportunity structures could fulfill these requirements, so the selection of a precise mechanism is to some extent arbitrary. I choose control of opportunity structures to represent intragroup control processes because it facilitates articulation with the two formal models upon which this paper is based (Heckathorn 1988, 1989), and because formal modelers have often neglected opportunity structures in favor of incentive structures, although both are obviously important.

Decision-making in the formal model is structured as follows: Each actor makes choices from among a specified group of alternatives that comprise the choice set. The choice set is assumed to consist of at most six alternative strategies. These include the four strategies defined above, and two oppositional strategies defined below. The opportunity of an actor to select any option $x$, $O_x$, is defined as the probability that this option is a member of the actor's choice set. For example, $O_d$ refers to the opportunity to defect at level one, and $O_c$ to the opportunity to cooperate at level two. Thus, if a rational actor attempts to defect at level one, he or she would have an $O_d$ probability of succeeding in that choice. The actor defects. On the other hand, the actor has a $1 - O_d$ probability of lacking the opportunity to defect. In that case, the actor is an involuntary cooperator, since he or she cooperates not through choice, but through lack of the opportunity to do otherwise. Finally, intragroup control is viewed as
reflecting changes in actors’ opportunity structures. Specifically, the efficacy, \( E \), of one actor’s control over another, is defined as the proportional change in the targeted actor’s opportunity to choose the targeted act. For example, \( E_2 \) refers to the efficacy of second level cooperation (i.e., compliant control), and \( E_3 \) to the efficacy of third level cooperation (i.e., oppositional control). As defined, the efficacy of control can vary from 1, where the other’s opportunity to select the option can be eliminated, to 0, where the other’s opportunity structure is unaffected.

To simplify further the formal model of the intragroup control process, exogenous limitations in opportunities are ignored. I assume that people begin with the ability to do anything they want. That is, opportunities to defect or cooperate at all three levels are unlimited (i.e., \( O_{d1} = O_{c1} = O_{d2} = O_{c2} = O_{d3} = O_{c3} = 1 \)). Intragroup controls operate by constraining their opportunities to do what they want. Specifically, they are assumed to use control at the second level only to promote cooperation at the first, so I treat that control as reducing \( O_{d1} \), while \( O_{d1} \) the opportunity to cooperate at level one, remains unrestricted. Similarly, when I consider oppositional control, it will be assumed that it reduces the opportunity to exercise compliant control, and that the opportunity to refrain from compliant control is always unrestricted. Thus intragroup control involves the restriction of only two opportunities, \( O_{d1} \) and \( O_{c2} \). In addition, I assume that the efficacy of control is uniform within the group. Thus, the group is initially assumed to be unstratified. (For an analysis of stratified systems using this model, see Heckathorn 1990a.)

The Risk of Sanctions

In the proposed model of group sanctioning, each actor’s risk of individual and collective sanctions is affected by his or her choice among the strategies of full cooperation, hypocritical cooperation, private cooperation, and full defection. Intuitively, it is clear that full defection maximizes that risk, whereas full cooperation minimizes it. The following integrates formal models of collective sanctions (Heckathorn 1988) and multilevel cooperation (Heckathorn 1989) to specify these risks quantitatively.

In systems that include collective sanctions, group members are jointly responsible for one another’s conduct. Sanctions are escaped if: (1) all group members conform voluntarily; (2) all potential violators lack the opportunity to consummate a violation; or (3) all violations escape detection. The risk that at least one potential violator will cause the group to be sanctioned, \( R_c \), can be straightforwardly computed. Consider first the prospect that collective sanctions are triggered by any single actor other than ego. That requires the conjunction of two events. Actor \( i \) must possess the opportunity to commit a first level defection, an event with probability \( O_{d1_i} \). Second, the agent must monitor that violation, an event with probability \( M_i \). Therefore, the risk of \( i \) committing a violation that triggers sanctions, \( R_i \), is the product of \( O_{d1_i} \) and \( M_i \), i.e.,

\[
R_i = O_{d1_i} M_i
\]

The probability that actor \( i \) will avoid triggering sanctions is thus \( 1 - R_i \). Furthermore, where the number of actors (other than ego) choosing to defect at the first level is \( N_{d1} \), the probability that all these actors (indexed as actors \( a, b, c \), through \( N_{d1} \)) will avoid triggering sanctions is the product of these probabilities, i.e., \( (1 - R_{a_i}) (1 - R_{b_i}) (1 - R_{c_i}) \ldots (1 - R_{N_{d1_i}}) \). Therefore, the risk that at least one potential violator will cause the group to be sanctioned, \( R \), is

\[
R = 1 - (1 - R_{a_i}) (1 - R_{b_i}) (1 - R_{c_i}) \ldots (1 - R_{N_{d1_i}})
\]

In expanded form,

\[
R = 1 - (1 - O_{d1_a} M_a) (1 - O_{d1_b} M_b) \ldots (1 - O_{d1_i} M_i)
\]

This is the risk of sanctions that our referent actor, ego, faces owing to the potential violations of other group members.

In a homogeneous group in which opportunities and monitoring are equivalent for all actors, equation 3 reduces to

\[
R = 1 - (1 - O_{d1} M)^{N_{d1}}
\]

This equation is subject to a second interpretation that does not require the somewhat counterfactual assumption of group homogeneity. The term \( (1 - O_{d1} M) \) can be interpreted as the geometric mean of actors’ probabilities of apparently conforming to the agent’s dictates. This in turn is related to the proportion of actors in the group who possess the opportunity to com-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
mit violations (O_1) and the proportion of violations that would be detected (M). R then describes the risk that a group member other than ego will trigger sanctions.

The risk of sanctions if ego chooses the strategy of full defection, R_{dd}, can now be computed. Avoiding sanctions in the case of double defection requires the conjunction of two events: first, ego must either lack the opportunity to commit a violation, or the violation must remain undetected, events with a combined probability of (1 - O_{dd} M). Second, other group members must fail to trigger sanctions (an event with probability 1 - R). Therefore, the probability of sanctions if the actor chooses double defection, R_{dd} is:

\[ R_{dd} = 1 - (1 - O_{dd} M) (1 - R) \]  
(5)

or, expanded and rearranged from equation 4,

\[ R_{dd} = 1 - (1 - O_{dd} M)^{N_{dd} - 1} \]  
(6)

Alternatively, through adopting the strategy of personal compliance, ego can reduce the risk of sanctions, R_{cd}, since then it is only the violations of others that subject the group to the risk of sanctions. Thus, if ego obeys the agent but free rides at the second (normative) level, the risk of sanctions that results if the actor chooses personal compliance, R_{cd} is merely R, i.e., from equation 4,

\[ R_{cd} = 1 - (1 - O_{dd} M)^{N_{cd}} \]  
(7)

Computing the risk of sanctions resulting from the exercise of normative control is somewhat more complex, since it requires that the model of intragroup control be further specified. Obviously, the creation of compliance norms yields benefits to the extent that it reduces the number of violations that risk triggering sanctions. Let E_{cd} be the efficacy of ego’s compliant control over actor i, defined formally as the proportional reduction in i’s opportunity to defect at the first level due to ego’s exercise of second level control, and let O_{cd} be ego’s opportunity to exercise that control. Thus, where O_{di} is i’s opportunity to defect at the first level prior to ego’s exercise of normative control, then i’s opportunity to defect subsequent to control, O’_{di}, is:

\[ O'_{di} = O_{di} (1 - O_{cd} E_{cd}) \]  
(8)

For example, if i initially has an 80 percent opportunity to defect at the first level, and ego’s efficacy of control is 30 percent with an unlimited opportunity to exercise that control, i’s opportunity to defect subsequent to control falls to 0.8 (1 - 0.3 • 1) = 56 percent. Ego’s opportunity to defect is assumed to remain unchanged, i.e., I assume that compliant control does not entail control of self.

Ego may not be the only actor who exercises control (a point which becomes particularly significant when a dynamic model is introduced). For example, if an actor i is controlled by N_{cd} actors other than ego (i.e., the number of actors choosing second level cooperation), and if only that control impedes i’s opportunity to defect, it follows from recursive application of equation 8 that the first actor (a) reduces the opportunity to defect to (1 - O_{cd} E_{cd}), the second actor (b) further reduces it to (1 - O_{cd} E_{cd} E_{cd}), and the N_{cd}th actor reduces it to

\[ O_{di} = (1 - O_{cd} E_{cd} E_{cd} \ldots (1 - O_{cd} E_{cd} E_{cd}) \ldots (1 - O_{cd} E_{cd} E_{cd}) \ldots ) \]  
(9)

If intragroup control is uniform across actors, this expression can be simplified:

\[ O_{di} = (1 - O_{cd} E_{cd})^{N_{cd}} \]  
(10)

If ego then contributes to these control efforts, that increases the number of second level cooperators from N_{cd} to N_{cd} + 1, so the opportunities to defect at level one subsequent to ego’s control, O’_{di}, are:

\[ O'_{di} = (1 - O_{cd} E_{cd})^{N_{cd} + 1} \]  
(11)

For example, if there are N_{cd} = 6 second level cooperators (other than ego) who possess unrestrained opportunities to cooperate at that level (O_{cd} = 1), and the efficacy of compliant control is E_{cd} = 0.2, the violation opportunity is reduced from a precontrol level of O_{di} = (1 - 0.2)^6 = 26 percent to a postcontrol level of O’_{di} = (1 - 0.2)^{6+1} = 21 percent.

Encouraging others to conform with the agent’s dictates in a mixed sanction system (i.e., cooperation at level two) is one way an individual can reduce the probability of group sanctions. This is the case regardless of the actor’s own decision to cooperate or defect at level one. The risk of collective sanctions if ego cooperates at the first and second levels, R_{cc}.
by both obeying the agent's dictates and encouraging others to do the same, is given by substituting the postcontrol violation opportunity $O'_{dl}$ into equation 7's expression for $R_{cd}$:

$$R_{cc} = 1 - (1 - O'_{dl}) M^{N_{dl}}$$  \hspace{1cm} (12)

or in expanded form:

$$R_{cc} = 1 - [1 - (1 - O'_{c2} E_{c2})^{N_{c2} + 1}] M^{N_{dl} + 1}$$  \hspace{1cm} (13)

According to this expression, the probability of sanctions declines to 0 if ego's control is complete and his or her opportunities to control are unrestrained (i.e., if $E_{c} = 1$ and $O'_{c2} = 1$, then $R_{cc} = 0$), for ego can then block any violation. If ego's ability to exert control is smaller, the probability of sanctions increases correspondingly, despite ego's own full compliance.

The probability of group sanctions if ego violates the agent's dictates but participates in sanctioning can be derived in a similar manner. Merely substitute $O'_{dl}$ for $O_{dl}$ in equation 6. The probability of group sanctions resulting from hypocritical cooperation, $R_{dc}$, is:

$$R_{dc} = 1 - (1 - O'_{dl}) M^{N_{dl} + 1}$$  \hspace{1cm} (14)

and in expanded form,

$$R_{dc} = 1 - [1 - (1 - O'_{c2} E_{c2})^{N_{c2} + 1}] M^{N_{dl} + 1}$$  \hspace{1cm} (15)

If ego's control over the group is complete and unconstrained, the probability of sanctions reduces to the prospect that ego's sole violation is detected (i.e., if $E_{c} = 1$ and $O'_{c2} = 1$, $R_{dc} = M$). If ego's control is less complete or more restricted, ego's own deflection constitutes but one possibly detected violation in the group (i.e., $R_{dc} > M$).

Payoffs from First and Second Level Cooperation

The probability of sanctions in a mixed sanction system is maximized by double deflection, minimized by double cooperation, and some intermediate probability of sanctions results from partial cooperation (i.e., hypocritical or private cooperation). That is,

$$R_{dd} \geq (R_{dc} \text{ or } R_{cd}) \geq R_{cc}$$  \hspace{1cm} (16)

To maximize the chances of gaining rewards or avoiding punishment, full cooperation is obviously the best choice. But full cooperation entails the costs of first and second level cooperation, and hence may be less than optimal.

The actor's payoff from choosing full cooperation in a mixed sanction system is a function of three interrelated factors. One is the prospect that another actor's violation will trigger collective sanctions. The expected value of this factor depends on both the strength of the collective sanction, $S_{c}$, and the risk that it will be administered given the actor's full cooperation, $R_{cc}$, or $-S_{c} R_{cc}$. In addition, the actor bears both the cost of first level cooperation, $K_{1}$, and the cost of second level (compliant) cooperation, $K_{2}$. The latter however, is weighted according to the probability that the actor will have the opportunity to exercise compliant control, $O_{c2}$, so that the expected value of second level cooperation is $-K_{2} O_{c2}$. Assuming von Neumann-Morgenstern utility scales, the payoff from full cooperation, $P_{cc}$, is the sum of the above three expected values:

$$P_{cc} = -S_{c} R_{cc} - K_{1} - O_{c2} K_{2}$$  \hspace{1cm} (17)

The payoff from full cooperation is a decreasing function of the strength of the collective sanction ($S_{c}$), and the costs of first and second level cooperation ($K_{1}, K_{2}$). However, determining when full cooperation is the optimal choice for a rational actor requires considering its payoff compared to the other possible strategies that involve either partial cooperation or no cooperation at all.

Similarly, one may calculate the payoff from private cooperation by combining the costs of the collective sanction and first level cooperation,

$$P_{cd} = -S_{c} R_{cd} - K_{1}$$  \hspace{1cm} (18)

The payoff from hypocritical cooperation, $P_{dc}$, combines the costs of the collective sanction, the individual sanction (weighted according to the probability of having the opportunity to commit a detected violation, $M O_{dl}$), involuntary first level cooperation [i.e., since the actor had chosen to defect at level one but may lack the opportunity to do so, the cost of first level cooperation is weighted according to the probability that the actor lacks the opportunity to defect, or $-K_{1} (1 - O_{dl})$, and second level cooperation,
decision can be changed only during his or her decision period. However, an actor’s behavior can change from period to period, since changes in opportunities are assumed to have immediate effects. For example, an actor who had chosen first level defection becomes an involuntary cooperator if the opportunity to defect at that level is subsequently lost.

Each actor is assumed to choose the strategy that maximizes his or her payoff during that decision period. In this sense, the actors are assumed to behave rationally. Consistent with my focus on the emergence of cooperation, actors are assumed to begin in a state of universal full defection, i.e., $N_{s1} = N_{c1} = 0$. At various points below, the effects of relaxing these assumptions are discussed.

Figure 1 depicts the emergence of cooperation in a representative mixed sanction system where $S_1 = 925, S_2 = 100, K_1 = 100, K_2 = 3, E_1 = 0.2, M = 0.1$, and $N = 10$. In this system, the spill-over from the individual to the collective sanction is rather modest (i.e., $S_1/S_2 = 100/925 = 11$ percent), and intragroup control is cheap relative to the cost of first level cooperation. Cooperation arises in a relatively complex manner. The first actor chooses hypocritical cooperation, and all other actors are assumed to begin in a state of full defection. Actor one reduces the opportunities of the others to defect at level one from $O_{d1} = 1$ to $O_{d1} = (1 - E_1 O_{c2}/K_2)^{S_1/S_2} = (1 - 0.2 \times 1)^{100/925} \approx 80$ percent. Therefore, 20 percent of the other nine actors lose the opportunity to continue defecting, so on average, 1.8 become involuntary first level cooperators. Therefore, in this group of ten actors, the mean group compliance level increases during the first period to 18 percent. The second through eighth actors choose the same strategy during their corresponding periods, further increasing the group compliance level. This is consistent with my argument elsewhere (Heckathorn 1988, p. 97) that hypocritical cooperation plays a crucial initial role in the emergence of collective action. The ninth and tenth actors choose personal compliance because the comparatively small number of violators remaining in the system weakens their incentive to participate in normative control, while their incentive for compliance is strengthened since the marginal risk of triggering sanctions grows with the shrinking number of violators.

For a sequential decision model that shares these first two assumptions, see Oliver et al. (1985, p. 333). See also Coleman (1986a).
quent periods, an equilibrium level of group compliance is attained, with 40 percent private cooperation and 60 percent hypocritical cooperation. This result is a noncooperative Nash (1951) equilibrium, in that no actor can gain by unilaterally departing from these strategies. Consequently, this equilibrium is stable.

In Figure 1, forty iterations were computed and equilibrium is attained after a dozen periods. All sequential models discussed in this article behave in a similar fashion. Therefore, the group compliance level at equilibrium for a system can be defined quantitatively as the mean compliance level (i.e., the mean proportion of voluntary and involuntary first level cooperators) during the last half of the periods.

Figure 1’s mixed sanction system illustrates the possibility of a division of labor between primary and second level cooperation. Hypocritical cooperators bear the costs of the normative control that compel other group members to bear the costs of primary level cooperation. Thus, the hypocritical cooperators function as managers, whereas the other actors function as workers. This outcome resembles Williamson’s (1975) account of the rise of internal organization. In both cases, a control structure with role differentiation arises to prevent opportunism from blocking collective action.

Sanction Strength, Spill-Over, and Compliance

A recurrent conclusion from the social psychological and deterrence literature is that individual sanctions alone have a limited and erratic ability to affect behavior (Brier and Fienberg 1980). This is consistent with the model, as illustrated in Figure 2. The vertical axis represents group compliance level (i.e., the equilibrium proportion of actors who comply either voluntarily or involuntarily), and the horizontal axis represents the strength of the individual sanction. The four lines plotted in the graph depict the group compliance level resulting from different levels of spill-over ($S_i/S_{ij}$) from the individual to the collective sanction. The line labeled “0%” depicts the limiting case of pure individual sanction systems that are devoid of any collective sanction. The result is a simple step function, with either 0 or 100 percent compliance. In essence, deterrence fails unless the cost of first level cooperation ($K_i$) is less than the anticipated costs of first level defection ($S_i M$). Thus, deterrence fails in the absence of either suitably strong individual sanctions, or sufficiently effective monitoring.

In mixed sanction systems, the determinants of compliance are more complex, since individual and collective sanctions have interactive effects. Figure 2 depicts three sets of mixed sanction systems in which the strength of the collective sanction corresponds to 5, 10, or 50 percent of the individual sanction strength. The results are counterintuitive. Even the modest spill-over of 5 percent has notable effects on group compliance. For example, when compared to the 0 spill-over case, 20 percent weaker sanctions (i.e., $S_i = 800$, and $S_j = 800 \cdot 5\% = 40$) induce about half of the group to comply. This result illustrates the compliance amplification effect of collective sanctions. More extensive spill-over produces far stronger amplification effects. For example, when spill-over is 10 percent, a 50 percent weaker sanction (i.e., $S_j = 500$, and $S_i = 500 \cdot 10\% = 50$) prompts more than 60 percent of the group to comply, and if spill-over is increased to 50 percent, an 80 percent weaker sanction (i.e., $S_i = 200$, and $S_j = 200 \cdot 50\% = 100$) prompts more than 80 percent of the group to comply.

Figure 2 implies that groups with high degrees of material interdependence are potentially far more tractable than atomized groups. Thus, in nonhomogeneous groups, differential

Note: $K_i = 100, K_j = 3, E_i = 0.2$, and $M = 0.1$. This graph shows the group compliance level resulting from spill-over levels of 0, 5, 10, and 50%. In the pure individual sanction system, where spill-over is 0, the relation is a step function with 0 compliance below 1,000 and perfect compliance above that level. In the mixed sanction systems, where spill-over varies from 5% to 50%, the relation is more complex. A spill-over of even 5% has a significant effect in promoting cooperation, implying that even modest amounts of spill-over may boost cooperation levels.
responses to sanctions can be expected theoretically. These conclusions are consistent with Hirschi’s control theory (1969), which explains conformity as resulting from links to parents, peers, school, and related institutions. He views links as having four components: attachment, involvement, belief, and commitment, all of which are theoretically expected to be positively related to spill-over, since they all have the effect of bonding the group into an interdependent whole. Indeed, if spill-over is viewed as a theoretic specification of the concept of linkage, the proposed model serves as a partial formalization of control theory.

The strength of the sanction amplification effect is highly dependent on intragroup control costs ($K_z$). If these costs are too high, second level cooperation ceases to serve as a meaningful option and the system functions merely as an individual sanction system. The two independent components of intragroup control costs are discussed by Heckathorn (1988, p. 547-48). One is the resource cost of control, e.g., the costs of monitoring intragroup compliance, sanctions, or other resources expended during the control process, and opportunity costs. There are theoretic reasons to suppose that at least in small scale social groups, resource costs tend to be low because peer approval is an extraordinarily potent social sanction. For example, some studies of deterrence indicate that informal sanctions may have greater deterrence effects than legal sanctions (Anderson, Chiricos, and Waldo 1977). What is significant in the present context is that peer approval is an inexpensive sanction that is available in full measure even to the most impoverished groups. Similarly, as Hechter (1984, p. 175) notes, groups with a high exit cost always have an inexpensive sanction at their disposal - the threat of expulsion.

The second component of compliant control cost is the contingent cost of control (Heckathorn 1988, pp. 547-48), the cost resulting from preexisting regulatory interests in the group. If an act directly benefits or harms other group members, these externalities give the group an incentive to regulate that behavior irrespective of any external sanctions. For example, consider the structure of group regulatory interests regarding heavy use of a drug such as crack, versus its sale. In the former case, the regulatory interest of a family would be uniformly negative, given the risk of adverse physical reactions and reduction in the individual’s capacity to contribute to the group. Independent of any spill-over from external sanctions, intragroup control then serves the interests of group members, and thereby deducts from the costs of control. In contrast, the sale of drugs can generate substantial positive externalities by contributing to family income. Hence the group gains an incentive to promote the activity, and loss of these potential benefits adds to the costs of control. For example, according to an urban anthropologist (Williams 1989, p.10), parents in inner city neighborhoods sometimes endorse their children’s drug dealing because it contributes to household income.

If an agent’s dictates coincide with preexisting regulatory interests within the group, the effect is to reduce the contingent costs of compliant control, whereas if the dictates conflict with these regulatory interests, the costs of compliant control increase. The implication is that control costs will tend to be low if the group’s preexisting regulatory interests are either neutral or coincident with the agent’s dictates.

The strength of the sanction amplification effect is highly dependent on intragroup control costs ($K_z$). If these costs are too high, second level cooperation ceases to serve as a meaningful option and the system functions merely as an individual sanction system. The two independent components of intragroup control costs are discussed by Heckathorn (1988, p. 547-48). One is the resource cost of control, e.g., the costs of monitoring intragroup compliance, sanctions, or other resources expended during the control process, and opportunity costs. There are theoretic reasons to suppose that at least in small scale social groups, resource costs tend to be low because peer approval is an extraordinarily potent social sanction. For example, some studies of deterrence indicate that informal sanctions may have greater deterrence effects than legal sanctions (Anderson, Chiricos, and Waldo 1977). What is significant in the present context is that peer approval is an inexpensive sanction that is available in full measure even to the most impoverished groups. Similarly, as Hechter (1984, p. 175) notes, groups with a high exit cost always have an inexpensive sanction at their disposal - the threat of expulsion.

The second component of compliant control cost is the contingent cost of control (Heckathorn 1988, pp. 547-48), the cost resulting from preexisting regulatory interests in the group. If an act directly benefits or harms other group members, these externalities give the group an incentive to regulate that behavior irrespective of any external sanctions. For example, consider the structure of group regulatory interests regarding heavy use of a drug such as crack, versus its sale. In the former case, the regulatory interest of a family would be uniformly negative, given the risk of adverse physical reactions and reduction in the individual’s capacity to contribute to the group. Independent of any spill-over from external sanctions, intragroup control then serves the interests of group members, and thereby deducts from the costs of control. In contrast, the sale of drugs can generate substantial positive externalities by contributing to family income. Hence the group gains an incentive to promote the activity, and loss of these potential benefits adds to the costs of control. For example, according to an urban anthropologist (Williams 1989, p.10), parents in inner city neighborhoods sometimes endorse their children’s drug dealing because it contributes to household income.

If an agent’s dictates coincide with preexisting regulatory interests within the group, the effect is to reduce the contingent costs of compliant control, whereas if the dictates conflict with these regulatory interests, the costs of compliant control increase. The implication is that control costs will tend to be low if the group’s preexisting regulatory interests are either neutral or coincident with the agent’s dictates.

The strength of the sanction amplification effect is highly dependent on intragroup control costs ($K_z$). If these costs are too high, second level cooperation ceases to serve as a meaningful option and the system functions merely as an individual sanction system. The two independent components of intragroup control costs are discussed by Heckathorn (1988, p. 547-48). One is the resource cost of control, e.g., the costs of monitoring intragroup compliance, sanctions, or other resources expended during the control process, and opportunity costs. There are theoretic reasons to suppose that at least in small scale social groups, resource costs tend to be low because peer approval is an extraordinarily potent social sanction. For example, some studies of deterrence indicate that informal sanctions may have greater deterrence effects than legal sanctions (Anderson, Chiricos, and Waldo 1977). What is significant in the present context is that peer approval is an inexpensive sanction that is available in full measure even to the most impoverished groups. Similarly, as Hechter (1984, p. 175) notes, groups with a high exit cost always have an inexpensive sanction at their disposal - the threat of expulsion.

The second component of compliant control cost is the contingent cost of control (Heckathorn 1988, pp. 547-48), the cost resulting from preexisting regulatory interests in the group. If an act directly benefits or harms other group members, these externalities give the group an incentive to regulate that behavior irrespective of any external sanctions. For example, consider the structure of group regulatory interests regarding heavy use of a drug such as crack, versus its sale. In the former case, the regulatory interest of a family would be uniformly negative, given the risk of adverse physical reactions and reduction in the individual’s capacity to contribute to the group. Independent of any spill-over from external sanctions, intragroup control then serves the interests of group members, and thereby deducts from the costs of control. In contrast, the sale of drugs can generate substantial positive externalities by contributing to family income. Hence the group gains an incentive to promote the activity, and loss of these potential benefits adds to the costs of control. For example, according to an urban anthropologist (Williams 1989, p.10), parents in inner city neighborhoods sometimes endorse their children’s drug dealing because it contributes to household income.

If an agent’s dictates coincide with preexisting regulatory interests within the group, the effect is to reduce the contingent costs of compliant control, whereas if the dictates conflict with these regulatory interests, the costs of compliant control increase. The implication is that control costs will tend to be low if the group’s preexisting regulatory interests are either neutral or coincident with the agent’s dictates.
effect is also affected by the agent’s monitoring capacity \((M)\) in complex and counterintuitive ways illustrated in Figure 3. Starting from a 0 level of monitoring, compliance increases monotonically, collapses to 0, and increases again. These relations become explicable when the joint effects of individual and collective sanctions are considered.

A previous study of collective sanctions (Heckathorn 1988, p. 557-58) showed that the relationship between monitoring and compliance in these systems is nonmonotonic. If there is too little monitoring, the agent is ignored, and if there is too much monitoring the group may give up on compliance on the expectation that someone will always trigger sanctions. Hence, group compliance can peak at intermediate levels of monitoring. In Figure 3’s system, compliant control (the dashed line) peaks at a 0.2 monitoring level, producing a localized increase in the compliance level (the fine line). The delinquency literature provides examples of such an anomalous relation between control and compliance. Wells and Rankin (1988, p. 280) found that too much or too little parental control is associated with greater frequency of delinquent behavior. Similar results were found in studies of adolescent deviance (Miller, McCoy, Olson, and Wallace, 1986). These results become explicable in terms of the proposed model if parental control of adolescents is viewed as a form of monitoring whose effects are mediated through peer group norms.

In contrast, the relationship between individual sanctions and monitoring is straightforward, at least in so far as actors behave rationally. In models of individual deterrence, monitoring is positively related to compliance. In Figure 3, above the 65 percent monitoring level, the level of compliant control (dashed line) is 0, but the compliance level (fine line) is high and results solely from individual deterrence. This is not to suggest that sufficiently high levels of monitoring will always yield high compliance levels in mixed sanction systems. If the individual sanction is weaker than the cost of compliance (i.e., if \(S < K\)), even perfect monitoring fails to produce individual deterrence.

In actual systems, exogenous constraints may place an upper limit on monitoring capacity that prevents individual deterrence from ever becoming effective. For example, most parents have only a limited opportunity to monitor the behavior of their adolescent children. Consequently, delinquency theorists such as Hirschi (1969) have dismissed direct parental control as a significant determinant of delinquency and focus on indirect controls based on the strength of ties to parents, peers, and school.

Finally, group size is an important determinant of the strength of the sanction amplification effect. As in the case of monitoring effects, it is useful to distinguish between the impact of collective and individual sanctions. In collective sanction systems, compliance incentives are greatest in moderately sized groups where optimal group size is a function of the strength of the collective sanction, the agent’s monitoring capacities, the efficacy of compliant control, and the proportion of first level defectors in the group. In very large groups, the compliance-inducing effects of collective sanctions disappear (Heckathorn 1988, pp. 550-52). Therefore, the relationship is nonlinear. In contrast, in individual sanction systems, group size has no independent effect except in so far as it is linked to otherwise significant parameters such as the agent’s monitoring capacity, which tends to decline with increases in group size. Therefore, the theoretic effects of group size in mixed sanction systems essentially re-
duce to those found in collective sanction systems.

In sum, the sanction amplification effect helps to resolve what may be termed the deterrence puzzle. Given the well-documented fact that individual deterrence does not work or works poorly, why is there not vastly more deviance and crime? One traditional answer to this variant of the Hobbesian problem of order is to appeal to the internalization of norms to explain the anomalous force of norms and laws. Yet that approach is dubious given the prevalence of opportunistic behavior. A more promising approach is to see whether group dynamics that bolster social control have been overlooked. Collective sanctions that generate compliance norms provide just such a dynamic by serving to reinforce intragroup normative constraints.

The proposed model suggests that the traditional analytic emphasis on individual sanctions in studies of social and organizational control may be incorrect. At least in systems where actors are interdependent and possessed of internal control capacity, spillover effects from individual to collective sanctions may account for most of the compliance resulting from social sanctions.

OPPOSITION NORMS AND INTRAGROUP CONTROL

It is now almost commonplace to view social norms as springing up to resolve prisoner’s dilemma-type collective action problems, as in Hardin’s (1968) classic essay, “The Tragedy of the Commons.” The previous section’s analysis of the emergence of compliance norms is an example. However, the opposite situation is also possible. Consider the Tragedy of the Lawns: each member of a suburban community prefers that neighbors spray, clip, and edge their lawns until they reach perfection. As a result, the norms for lawn care become ever more stringent, extending far beyond simple mowing and edging to expensive and time-consuming procedures at the forefront of the science of lawn care. As the burden of lawn care grows, residents yearn nostalgically for the days of simple mowing.

Rather than resolving collective action problems, social norms can create them by mandating compliance even when the individual costs exceed the collective gain. Resolving this prisoner’s dilemma requires overcoming such norms — the collective action problem is turned on its head. To defect is to participate in norm creation and enforcement, and to cooperate is to let others do as they wish. This situation may be termed an inverted prisoner’s dilemma. When a group creates norms in the face of this dilemma, the outcome can be termed over-control, and if these norms are obeyed, the Pareto-deficient outcome is an example of over-compliance.

Over-control can occur in any type of behavioral control system if compliance persists despite increasing compliance costs (e.g., the speed-up of an assembly line) or weakening external sanctions (e.g., pay cuts). (See Coleman 1983 for a related analysis.) For example, consider the effects of changing Figure 2’s collective sanction system so that the strength of the individual sanction, $S_i$, is set at 400, the collective sanction, $S_c$, is set at 200, and the cost of first-level cooperation, $K_i$, is vastly increased to 1,000. Using a monetary example, this corresponds to a case where an agent threatens group members with individual fines of $400 and group fines of $200 (per member) if any group member fails to expend $1,000 complying with the agent’s dictates. It would be absurd for the group to comply since compliance is far more costly than the threatened sanctions, i.e., $K_i > S_i + S_c$, since $1,000 > 400 + 200$. Yet this is not the end of the story. In this system, each actor prefers that other actors comply. Only in that way do others act to reduce the chances of collective sanctions (e.g., if 10 percent of noncompliant acts are detected by the agent, the first violation subjects the group to a sanction risk of 10 percent, a second increases that risk by another 9 percent, and so forth). Thus, the individual acquires a regulatory interest or preference that others comply. Under certain circumstances, these regulatory interests can serve as the basis for a group norm that compels extraordinarily costly compliance. In that case, rather weak regulatory interests triumph over far stronger contrary inclinations.

If the formal model is applied to the above example, the result is over-compliance, as illustrated by the bold line in Figure 4. Despite the collective and individual irrationality of compliance within this system, the equilibrium compliance rate reaches .89. This occurs not because any actors choose first level cooperation, but because the universal choice of hypocritical cooperation coheres compliance. Hypo-
critical cooperation constitutes defection in an inverted prisoner’s dilemma. In essence, the relatively weak regulatory interest in promoting compliance by others creates compliance norms that defeat the stronger individual interests in noncompliance.

Over-compliance of the sort illustrated in Figure 4 is probably not rare empirically. Many social groups are characterized by an instrumentally excessive emphasis on conformity. As Coleman (1987, p. 148) observes, some norms regulate behaviors that appear to lack externalities. In that case, any normative control is over-compliance. Similarly, anomalously high levels of social cooperation underlie the paradox of voting, and the Hobbesian problem of order. According to some critics, an exaggerated concern with conformity pervades much of American middle class culture. However, over-compliance is obviously not universal. Countervailing mechanisms arise to block the emergence of excessive compliance norms or to reduce their efficacy.

Oppositional Control

Thus far, social actors have been assumed to face a twofold choice at the normative level — either they cooperate by supporting compliance norms, exercising compliant control, or they defect and leave others to behave as they wish. Control can also be directed toward other objectives, including opposing the control attempts of other actors. Whereas compliant control is directed at affecting the first level behavior of other actors, oppositional control is directed at second level behavior. Thus, it is a “meta-norm” (Axelrod 1986), a norm that regulates a norm, and constitutes a form of third level cooperation that can give rise to a third level freerider problem.

With the inclusion of oppositional control as a third level strategy, actors acquire additional strategic choices. If one ignores the possibility that an actor chooses the self-defeating strategy of exercising both compliant and oppositional control, there are two new strategic options. In addition to the strategies of full cooperation, hypocritical cooperation, private cooperation, and full defection, they also can choose either hypocritical opposition, i.e., cooperating at the first level while exercising oppositional control at the third, or full opposition, i.e., defecting at the first level while exercising oppositional control. This section explores the optimality conditions for the exercise of oppositional control, with special emphasis on the

$$N_0 = 0, N_{-1} = 9$$ is $-895$; and the payoff from unilateral cooperation (i.e., full defection in a system where others choose the hypocritical defection strategy, $P_{48}$ if $N_{-1} = N_{0} = 0$) is $-896$. These payoffs have the specified order. The second defining characteristic of a prisoner’s dilemma is that a mixture of unilateral cooperation and unilateral defection yields a lower payoff than does universal cooperation. This condition is satisfied in Figure 4 because the mixture of payoffs of -156 and -896 is lower than a payoff of -170. Therefore, Figure 4 portrays a prisoner’s dilemma. It differs from other prisoner’s dilemmas only in that the meanings of the terms “cooperate” and “defect” are inverted from their usual meanings.
opposition between compliant and oppositional control.

Oppositional control can be modeled in the same manner as compliant control. The opportunity of group members to exercise compliant control, \( Q_{c2} \), can be represented as a function of both the efficacy of oppositional control, \( E_{c3} \), and the number of actors exercising that control, \( N_{c3} \):

\[
O_{c2} = (1 - E_{c3})^{N_{c3}}
\]  

(21)

This expression summarizes the total effect of oppositional control exercised by actors other than ego upon opportunities for group members to support compliance norms. If ego also exercises oppositional control, that increases the number of actors in opposition from \( N_{c3} \) to \( N_{c3} + 1 \), further reducing opportunities to support compliance norms. Therefore, actors’ opportunities to cooperate at the second level resulting from ego’s oppositional control are:

\[
O''_{c2} = (1 - E_{c3})^{N_{c3} + 1}
\]  

(22)

According to this expression, opportunities to support compliance norms decline to 0 if ego’s oppositional control is total (i.e., if \( E_{c3} = 1 \), then \( O''_{c2} = 0 \)). If ego’s oppositional control is weaker, opportunities to support oppositional norms are correspondingly greater.

Exercising oppositional control yields benefits to ego only to the extent that it expands opportunities to avoid costly compliance, i.e., to the extent that it increases the opportunity to defect at the first level (\( O_{d1} \)). An expression for the opportunity to defect at the first level after oppositional control, \( O''_{d1} \), is given by substituting the opportunity for compliant control as altered by ego’s oppositional control, \( O''_{c2} \), into equation 10’s expression for \( O_{d1} \):

\[
O''_{d1} = (1 - O''_{c2} E_{c3})
\]  

(23)

Oppositional control has impacts beyond changing the first and second level opportunities available to group members. It also affects the risk of the group being sanctioned. Consider first the risk of sanctions resulting from hypocritical opposition, \( R_{co*} \), a strategy of complying at the first level while exercising opposition at the higher level. Given that the actor complies at the personal level while choosing third rather than second level cooperation, the probability of sanctions that results if the actor chooses hypocritical opposition, \( R_{co*} \), is given by substituting \( O''_{d1} \) for \( O_{d1} \) in equation 12:

\[
R_{co} = 1 - (1 - O''_{d1} M)^{N_{d1}}
\]  

(24)

Alternatively, if the actor chooses the strategy of full opposition by combining opposition with personal noncompliance, the chances of group sanctions increase. It increases the number of actors choosing noncompliance from \( N_{d1} \) to \( N_{d1} + 1 \). Thus, the probability of sanctions that results if the actor chooses full opposition, \( R_{do*} \), is derived by substituting \( N_{d1} + 1 \) for \( N_{d1} \) in equation 24’s expression for \( R_{co*} \):

\[
R_{do} = 1 - (1 - O''_{d1} M)^{(N_{d1} + 1)}
\]  

(25)

Payoffs from Full and Hypocritical Opposition

The payoff to an actor choosing full opposition depends on four factors: (1) the expected cost of the collective sanction, \( S_{re*} \); (2) the expected cost of the individual sanction, \( S_{O''_{d1}} M \); (3) the cost of involuntary compliance, \( K_{1} (1 - O''_{c3}) \); and finally the cost of oppositional control, \( K_{3} \). The payoff resulting from full opposition combines these four terms:

\[
P_{do} = -S_{c} R_{do} - S_{1} M O''_{d1} - K_{1} (1 - O''_{d1}) - K_{3}
\]  

(26)

Similarly, the payoff resulting from hypocritical opposition combines the costs of the collective sanction and first and third level cooperation,

\[
P_{co} = -S_{c} R_{co} - K_{1} - K_{3}
\]  

(27)

Hypocritical opposition is necessarily suboptimal. The gains to an actor who chooses oppositional control derive exclusively from that control’s ability to reduce involuntary compliance. Yet hypocritical opposition involves choosing to comply, thus guaranteeing that the individual will bear the full costs of compliance. Consequently the actor fails to benefit from his or her own oppositional control.\(^{6}\)

---

\(^{6}\) Hypocritical opposition is not always suboptimal in more complex systems. For example, if the actor had a comparative advantage in exerting oppositional control (e.g., a high oppositional control level, \( E_{c3} \), or a low oppositional control cost, \( K_{3} \)), and
A Procedural Decision Model of Oppositional Control

The problem of over-compliance illustrated by Figure 4 is resolved if the group's strategic possibilities include oppositional control. When oppositional control costs are low (K_o = 3), oppositional control emerges very quickly. The first actor chooses hypocritical cooperation, but subsequent actors choose full opposition, so the equilibrium level of compliance falls to only 3 percent. By exercising oppositional control, group members beat down the compliance level to a collectively more rational level. The victims of the "tragedy of the lawns" may therefore fight for the right not to edge their yards. Similarly, juveniles may fight against conventional values in the manner described in Cohen's (1955) account of the emergence of delinquent subcultures. From the standpoint of juveniles facing failure at school and grim prospects in the labor market, adherence to conventional values may constitute a form of what is here termed over-compliance.

Figure 4 implies that increasing oppositional control costs can serve to increase group compliance. If those costs are sufficiently high, group compliance levels can far exceed the point at which over-control begins. A number of techniques of group control can be understood in this way, including rules requiring group members to display a certain esprit de corps, organizational loyalty, professionalism, or "sprit." Religious sects frequently demand extensive compliance (Hechter 1987; Iannaccone 1988). These typically include rewards for compliant control (e.g., attempts to convert nonbelievers or reconvert lapsed followers) and punishments for oppositional control (e.g., fostering doubts in others). Thus, the costs of compliant control are reduced and the costs of oppositional control increase, which strengthens tendencies toward group over-compliance.

Since they require direct monitoring of intragroup control processes, prohibitions on oppositional control are not feasible in many social settings. Monitoring is possible only in special circumstances, such as co-residential religious communities where privacy is minimal.

However, other means are available for altering a group's propensity to engage in oppositional control. Reducing the efficacy of oppositional control (E) diminishes oppositional incentives. This can be accomplished through institutional measures that atomize the group, e.g., promoting mobility into and out of the group, transferring personnel within the group, and minimizing opportunities for extended personal interaction. However, such measures tend to reduce the group's internal control capacity and depress compliant control (i.e., both E and E are reduced). Hence, this approach can backfire by reducing group compliance levels.

The theoretic relationship between the efficacy of intragroup control (E and E) and compliance levels is surprisingly complex. If compliance is extremely costly relative to sanction strength, as in Figure 4, virtually any internal control capacity is channeled into oppositional control. Alternatively, if compliance costs are low, all control capacity is channeled into compliant control regardless of the efficacy of oppositional control. In intermediate cases, where sanction strength is moderate relative to compliance costs, the relationship is more interesting. Figure 5 depicts the relationship between efficacy of control and compliance level in Figure 2's system. Here, compliance level is an increasing function of efficacy, up to the 75 percent efficacy level. Above that point, the compliance level drops precipitously. If the group's cohesion exceeds a threshold value, the compliance level plummets.

The nonmonotonic relationship between group cohesion and compliance has implications for organizational control. Exertions of power frequently involve measures designed to atomize the target of control. The durability of resource costs of control, K_r, and the contingent costs of that control, K:

\[ K_r = K_{r3} + K_{r3} \]  
where \( K_{r3} \) is defined in a manner parallel to \( K_{r3} \), i.e.,

\[ K_{r3} = I_{r1} N_{r1} (O_{r1} - O'_{r1}) \]
promote compliance norms and bars foster oppositional norms. The Nazis’ consolidation of power in the mid-1930s provides a more extreme example. According to Allen (1965, pp. 209-26), the Nazis adopted a policy of promoting atomization and cohesion — they destroyed or reorganized communal endeavors, even chess clubs, so that all informal associations were directly under party control.

The impact of individual sanctions on opposition norms is worthy of note. Individual sanctions reduce incentives for groups to evolve oppositional norms, even if they are too weak to produce any individual deterrence. This effect is particularly important when the group’s intragroup control is high. In Figure 5, where control above a threshold value provoked opposition, the strengthening of individual sanctions blocked that revolt. The strength of the individual sanction required is positively related to the level of group control (e.g., at a level of 80 percent control, an increase in the individual sanction from $S_i = 100$ to $105$ is required to block opposition), whereas if the efficacy of control increases (e.g., to the 90 percent or 100 percent levels), only a moderately stronger sanction is required (e.g., $S_i = 110$ and 119, respectively). What is important is that individual sanctions that are far too weak to produce deterrence (i.e., less than $S_i = 1,000$ in Figure 5) can nonetheless block the development of oppositional norms.

In sum, according to the proposed model, oppositional control can arise within groups as a solution to an inverted prisoner’s dilemma problem in which universal defection produces not a breakdown of social cooperation but a collectively irrational excess of social cooperation. The emergence of oppositional control within a group depends on a number of factors, including the costliness of compliance relative to the strength of sanctions, the costliness and efficacy of oppositional control, and the group’s cohesion.

CONCLUSION

It is necessary to consider the time frame within which decisions about social control are made. External agents such as states that deploy sanctions through a judicial system, and corporations that deploy sanctions through bureaucratic structures, tend to change strategies slowly. In the judicial system, years may lapse between arrest and conviction. The typical interval be-
tween changes in the law are even greater. In the corporate world, weeks or months frequently separate misdeeds from subsequent demotion or dismissal, and fundamental corporate policies can remain stable for months or years.

In contrast, normative control proceeds swiftly. Only fractions of a second may elapse between a faux pas and looks of disapproval. Under appropriate circumstances, as in Gilham’s (1982) account of Marine Corps recruit training, norms can arise almost instantaneously. When viewed from the accelerated time frame of normative control, legal and other formal controls appear to be governed by fixed strategies.

The differences in time scale between formal and normative social control have significant implications. People vary substantially in their temporal orientations, i.e., in their ability to delay gratification, their degree of present orientation, or their discounting rate (Schelling 1984). When viewed as rational actors, they vary in the length of the interval over which they choose to maximize their payoffs. Hirschi and Gottfredson (1986, 1987) argue that differences in temporal orientation explain the generality of deviance, i.e., the strong positive correlation among forms of deviance. They argue that types of deviance all stem from a tendency to seek immediate pleasure or gratification.

Given the wide disparity of time scales in the operation of formal and normative controls, the distribution of temporal orientations is a potentially important determinant of group behavior in systems that combine both types of control. For example, an actor in a mixed control system with a strong present orientation is insensitive to formal sanctions but highly sensitive to normative controls. Hence, that actor remains tractable to intragroup controls. Like many of the actors in the systems analyzed, the actor conforms not because of individual deterrence but because of the group’s compliance norms.

So long as the proportion of actors with a strong present orientation does not become so great that it deprives the group of a sufficient number of motivated second level cooperators, present-oriented actors do not significantly affect system dynamics. The intragroup control system is able to contain their greater personal tendencies toward noncompliance. For example, in rural communities where normative controls are effective, a handful of present-oriented individuals can be effectively controlled. In contrast, if all members of a group possess a strong present orientation, this weakens the individual and collective sanctions that could motivate personal compliance and the creation of compliance norms (i.e., it reduces $S_1$ and $S_2$). Hence the group as a whole may become noncompliant. Furthermore, temporal orientation may serve as a basis for differential association, given that members of any group have an incentive to escape from the control of norms based on regulatory interests that they do not share. Groups which are relatively homogeneous in temporal orientation can arise in otherwise atomizing urban settings, with each group’s degree of conventionality dependent on its temporal orientation. These speculations may account for the strength and consistency of the association between urbanization and criminality and deviance.

I have examined the complexities that arise when interdependent groups of actors are subjected to social sanctions. Compliance or opposition to control in groups can arise from more complex processes than are possible in the two-person case. For example, compliance that appears as simple individual deterrence, may arise through more circuitous routes. Group-mediated sanction amplification effects can produce compliance even when sanctions are far too weak to produce individual deterrence. Alternatively, noncompliance that appears to be a failure of individual deterrence may result instead from opposition norms. According to the analysis, whether members of a group deploy their internal control capacity to reinforce or oppose external control of the group depends quite sensitively on the costs of compliance, the costs of intragroup control, the efficacy of intragroup control, the strength of external sanctions, and the external agent’s monitoring capacities.

Douglas Heckathorn is Professor of Sociology and Economics at the University of Missouri-Kansas City. This article grows out of a project focusing on links between normative and formal controls in large and small scale groups. His research interests also include work with Steven Maser and Jules Coleman to analyze bargaining and transaction costs in social and economic institutions. Recent articles from this project appear in American Journal of Political Science, Journal of Law, Economics, and Organization, Journal of Politics (in press), and Harvard Journal of Law and Public Policy. The latter article won the Lon Fuller Prize in Jurisprudence.
APPENDIX

The following illustrates the procedures for computing the outcome of the first period of Figure 1’s mixed sanction system, with parameters \( S = 100, \ S_1 = 925, \ E = 0.2, \ E_1 = 0.2, \ M = 0.1, \ K = 100, \ K_1 = 3, \ K_2 = 3, \) and \( N = 10. \) (A more detailed description is available from the author (Heckathorn 1990b).)

Actor 1 moves during the first period, actor 2 moves during the second, and so forth until period 10 when actor 1 makes a second move. Computing actor 1 (ego’s) optimal strategy during the first period requires the following steps:

1. Identify strategy selections by other actors in the system during the previous \( N-1 \) periods. Since the group is assumed to begin in a state of universal full defection at period 1, \( N_{0(1)} = 9, \ N_{1(1)} = 0, \) and \( N_{2(1)} = 0. \)

2. Compute opportunities in the system as a function of strategy selections.
   
   \[ O_{0(1)} = (1 - 1 \cdot 0.2)^0 = 1 \]
   \[ O_{1(1)} = (1 - 1 \cdot 0.2)^1 = 0.8 \]
   \[ O_{2(1)} = (1 - 1 \cdot 0.2)^2 = 1 \]
   \[ O_{3(1)} = (1 - 0.2)^0 = 1 \]
   \[ O_{4(1)} = (1 - 0.2)^1 = 0.8 \]

3. Compute the risks of sanctions as a function of the above computed strategy selections and opportunities:
   
   \[ R_{0(1)} = 1 - (1 - 0.8 \cdot 0.1)^0 = 0.527838 \]
   \[ R_{1(1)} = 1 - (1 - 0.8 \cdot 0.1)^1 = 0.565611 \]
   \[ R_{2(1)} = 1 - (1 - 0.1 \cdot 0.1)^0 = 0.612579 \]
   \[ R_{3(1)} = 1 - (1 - 0.1 \cdot 0.1)^1 = 0.651321 \]
   \[ R_{4(1)} = 1 - (1 - 0.1 \cdot 0.1)^2 = 0.651321 \]

4. Compute payoffs from alternative strategies as a function of the above computed opportunities and sanction probabilities:
   
   \[ P_{0(1)} = 100 \cdot 0.527838 \cdot 1 - 0.1 \cdot 1 - 3 = -155.783 \]
   \[ P_{1(1)} = 100 \cdot 0.565611 \cdot 925 \cdot 0.1 \cdot 1 - 100 \cdot 0.1 \cdot 1 - 3 = -152.061 \]
   \[ P_{2(1)} = 100 \cdot 0.612579 \cdot 100 \cdot 3 = -161.257 \]
   \[ P_{3(1)} = 100 \cdot 0.651321 \cdot 925 \cdot 0.1 \cdot 1 - 100 \cdot 0.1 \cdot 1 - 3 = -160.632 \]
   \[ P_{4(1)} = 100 \cdot 0.651321 \cdot 925 \cdot 0.1 \cdot 1 - 100 \cdot 0.1 \cdot 1 = -157.632 \]

5. Finally, based on the above computed utilities, identify ego’s optimal strategy, and update strategy selections in the system accordingly. Ego’s optimal strategy is hypocrical cooperation since \( P_{4(1)} \) is maximal. Therefore ego defects at the first level and cooperates by choosing compliant control at the second level, so the strategy selections during period 2 are:

   \[ N_{0(2)} = 9 + 0 = 9 \]
   \[ N_{1(2)} = 0 + 1 = 1 \]
   \[ N_{2(2)} = 0 + 0 = 0 \]

The outcomes of subsequent periods are computed by repeating steps 2 through 3 above.

REFERENCES


Green, James R. 1980. The World of the Worker: