Noncompliance and the limits of coercion: the problematic enforcement of unpopular laws

Huckfeldt, Robert

Noncompliance And The Limits Of Coercion
The Problematic Enforcement Of Unpopular Laws

Robert Huckfeldt

ZUMA-Arbeitsbericht Nr. 89/08

Zentrum für Umfragen, Methoden und
Analysen e.V. (ZUMA)
Postfach 122155
D-6800 Mannheim 1

Seit Juli 1983 sind die ZUMA-Arbeitsberichte in zwei Reihen aufgeteilt:

Die **ZUMA-Arbeitsberichte** (neue Folge) haben eine hausinterne Begutachtung durchlaufen und werden vom Geschäftsführenden Direktor zusammen mit den übrigen Wissenschaftlichen Leitern herausgegeben. Die Berichte dieser Reihe sind zur allgemeinen Weitergabe nach außen bestimmt.

Die **ZUMA-Technischen Berichte** dienen zur hausinternen Kommunikation bzw. zur Unterrichtung externer Kooperationspartner. Sie sind nicht zur allgemeinen Weitergabe bestimmt.

NONCOMPLIANCE AND THE LIMITS OF COERCION:
THE PROBLEMATIC ENFORCEMENT OF UNPOPULAR LAWS*

Robert Huckfeldt
Indiana University
May 1988

ABSTRACT

This paper constructs a mathematical representation of the dynamic logic tying together coercion and noncompliance within political systems. Several concepts are central to the effort: system legitimacy, the long-term institutionalized investment of the political system in coercive resources, the vulnerability of political systems to mass noncompliance, and the short-term response of the political system to noncompliance. The analysis addresses a number of issues. What are the dynamic consequences of system legitimacy and of the political system's institutionalized investment in coercive resources? What are the consequences of its short-term coercive response to noncompliance? Under what circumstances will noncompliance be eliminated? Under what circumstances will it be controlled? Under what circumstances will political systems become overwhelmed by noncompliance? A central argument of this paper is that aggressive short-term responses to noncompliance are likely to be ineffective in controlling noncompliance, and they are likely to aggravate the vulnerability of political systems to mass noncompliance.

*This research has been supported through the generosity of Hartmut Esser, Peter Mohler and other colleagues at Zentrum für Umfragen, Methoden und Analysen in Mannheim. I am also grateful to a number of colleagues who have provided helpful suggestions along the way: Robert Muncaster, Vincent Ostrom, Barbara Salert, Peter Schmidt, John Sprague, Harvey Starr, Robert Stein, several anonymous reviewers, and a number of perceptive graduate students.

Note: This paper is forthcoming in Mathematical Modelling.

NONCOMPLIANCE AND THE LIMITS OF COERCION:
THE PROBLEMATIC ENFORCEMENT OF UNPOPULAR LAWS

Robert Huckfeldt
Indiana University
May 1988

Some laws secure broad support and compliance within political systems because they are fundamental to the social order. Murder, extortion, theft, and kidnapping are broadly conceived as socially deviant behaviors, at least in nonrevolutionary circumstances, and the prohibition of these activities generates little controversy. In contrast, most political systems address an entirely different set of problems with respect to the enforcement of unpopular laws -- laws that proscribe activities <u>not</u> seen as deviant by large parts of the population. These laws are much more controversial because, in one way or another, the public disputes their validity. The enumeration of such laws is lengthy, and forbidden activities under various regimes range from (1) speeding on highways and smoking marijuana to (2) exercising free speech and participating in political assemblies. Unpopular laws are distinguished not by the seriousness of the activity they forbid, but rather by the level of public resistance they provoke. These laws generate an important political dimension because they involve the clashing interests of citizens and political systems. They produce a potential confrontation between the coercive power of the state and noncompliance on the part of the public.

This relationship between coercion and compliance is fundamental to politics, and to the viability of political systems. The maintenance of order and the survival of political systems require large scale compliance with politically established norms and legal mandates, and the state's coercive capacity is a central ingredient in securing compliance. Of course, citizens need not always be coerced, even to obey unpopular laws. Compliance is also secured voluntarily, absent even the threat of sanctions for noncompliance, and therein lies the central ingredient of system legitimacy. Furthermore, the coercive capacity of the state is limited, especially in democratic regimes, but also in political systems lacking popular control. At some level of noncompliance the coercive response of the state must reach its maximum and decline. In a democratic regime this is the very definition for popular control, and in a non-democratic regime it is the definition for revolution.

This paper represents the relationship between coercion and compliance in the form of a mathematical model. The model incorporates elements that are both dynamic and simultaneous in timing. The interdependence between coercion and compliance is viewed as a process that unfolds in time. Levels of coercion are preconditioned by previous levels of noncompliance, while levels of compliance are preconditioned by previous levels of coercion.

The mathematical representation is, in turn, employed to address several questions: Why do some efforts at enforcing unpopular laws fail, while others succeed? To what extent is success a function of system legitimacy and the response by citizens to sanctions for noncompliance? To what extent is success a function of the political system's long-term investment in coercive resources, and of its short-term coercive response to noncompliance? Under what circumstances can a political system obtain compliance with unpopular laws absent significant coercion? Finally, what are the strategic consequences of the limitations upon coercion as a strategy for securing compliance? How do these limitations affect the relationship between coercion and noncompliance, and the exercise of coercion by political systems?

A Characterization of Compliance

Why do people break laws? More importantly, why do some people quit obeying laws, and why do others quit breaking them? First, some people obey laws *because* they are laws, regardless of any penalties that might be levied against law breakers, and regardless of any benefits that might be obtained from noncompliance. Thus, there is some limit upon noncompliance under normal circumstances, and this limit is defined here as L: the proportion of citizens who *might* refuse to comply with a particular law (Likens and Kohfeld, 1983). *Legitimacy* is defined as the proportion who would never break the law under normal circumstances, 1-L.

Who then might become noncompliers? Let N_t denote the proportion of citizens who disobey the law at time t. It follows that the proportion of citizens eligible to become noncompliers at the next instant in time is equal to $L-N_t$, the proportion of citizens who might break the law, but who are currently obeying it. Thus L is defined as being fixed in time, while N_t varies through time. (In the dynamic representations of this paper, time subscripts are maintained on system states to emphasize the fact that they vary with time, and to call attention to the time invariant model parameters. All rates of change are instantaneous, however, and no time delays are built into the system.)

Compliers become noncompliers for two interdependent reasons: First, law abiding citizens are more willing to become lawbreakers to the extent that breaking a particular law is profitable, and the profitability of breaking a particular law is conceived as being constant in time (a). This is not to say that profitability is constant across crimes or across individuals. It is generally more profitable to break a speed law on a trip across Wyoming than it is to break a speed law on a trip to the grocery. Similarly it is more profitable for the employee who is late for work to break a speed law than it is for the employee who is on time. Thus the profitability of noncompliance (a) is defined as an average within a population relative to a particular law.

Profitability derives from numerous sources: some rioters steal color televisions from department stores in order to enjoy football games, while others enjoy the release of pent-up fury, while still others view their involvement as an instrumental act aimed at bringing down an oppressive political system. Indeed, economic analyses of compliance occupy a prominent place in the literature on crime and deterrence. A seminal work is that of Becker (1968), in which he argues that ". . . a useful theory of criminal behavior can dispense with special theories of anomie, psychological inadequacies, or inheritance of special traits and simply extend the economist's usual analysis of choice" (p. 170).

Profitability is not the only reason that law abiding citizens become law breakers, however. The individual level transformation from obeying to breaking the law is also more likely to occur if the general level of noncompliance is higher. Two different micro-level premises support this assertion. First, the rationally calculating, football-loving rioter might decide that it is safer to steal the television if many other people are stealing televisions, stereos and so forth. That is, the probability of individual apprehension decreases as the aggregate level of noncompliance increases (Granovetter, 1978). Thus, a social choice interpretation of individual behavior can be called upon to support the assertion that people are more likely to disobey a law to the extent that others also disobey the law.

Second, the importance of aggregate noncompliance for the individual decision to disobey a law might also be supported on the basis of a social structural interpretation of individual behavior. Some forms of noncompliance may be furthered through social interaction and social persuasion, and in these instances aggregate noncompliance is important as a source of social influence. At one extreme, revolutionary success requires revolutionary solidarity, and thus the revolution proceeds (diffuses) as revolutionaries are able to convert nonrevolutionaries. At some point, even many unwilling revolutionaries may be swept along by the social tide as the level of revolutionary activity increases. (For complementary analyses see Przeworski 1974, Tilly 1978, and DeNardo 1985: 45-46.)

Barry (1970) and Salert (1976) show that considerable difficulties occur when collective revolutionary activity is conceived in terms of individually based, maximizing behavior in the tradition of Olson (1971). These problems do not affect the present representation, however. The form of the model allows for either a social _choice_ interpretation, or a social _structure_ interpretation. One or the other will be more or less appropriate depending upon the form of noncompliance being considered. The important point is that both sets of micro-level premises support the importance of aggregate noncompliance as a factor affecting the individual's decision regarding compliance.

These two factors -- profitability and aggregate noncompliance -- are combined interactively as a product to form a rate at which compliers become noncompliers: aN_t. And thus the _increase_ in noncompliance is defined as: $aN_t(L - N_t)$.

Why do law breakers become law abiding citizens? They respond to the coercive sanctions of the state, either because they are apprehended and punished, or because they fear such apprehension. These coercive sanctions may usefully be conceived as two separate components: the resources available to the coercive apparatus, and the effectiveness of the coercive apparatus. For example, Tsar Nicholas I invested heavily in coercive resources to silence the Decembrists and other political critics, but the effort was undermined by the ineffectiveness of the coercive apparatus. The effectiveness of the state's coercive effort is determined by a variety of factors: corruption within the system, institutional inefficiency, constitutional safeguards for citizens, and in the case of Russia's Third Section, the assignment of mutually contradictory and impossible tasks (Monas, 1961). These factors tend to be deeply ingrained within particular political systems, and thus effectiveness is treated as being constant in time with respect to a particular law (b).

In contrast, coercive resources are highly variable across time. A typical short-term response to higher levels of noncompliance is additional investment in coercive resources. Thus, coercive resources are defined relative to time as P_t, and the rate at which noncompliers become compliers is defined as an interaction between these two factors, written as the simple product: bP_t. This means, in turn, that the <u>decrease</u> in noncompliance is represented as: bP_tN_t.

As the works of Gurr (1970: chap. 8) and Salert and Sprague (1980) demonstrate, the public is not always encouraged to comply by increased levels of coercion. Increased coercion is capable of provoking outrage rather than restraint, thereby producing increased levels of noncompliance. This possibility generates important analytic consequences, but the focus of the present effort is upon variations in the response of the political system to noncompliance, rather than variations in the response of citizens to coercion.

Exploring both types of variations simultaneously lies beyond the bounds of this paper, and the present effort must be understood subject to the assumption that noncompliance is discouraged by increased coercion.

In summary, the net rate of <u>change</u> in noncompliance relative to time (dN_t/dt) is represented as the increase in noncompliance minus the decrease in noncompliance, or:

$$dN_t/dt = aN_t(L-N_t) - bP_tN_t. \quad (1)$$

A Characterization of Coercion

What factors give rise to changes in the resource level devoted by the political system to coercion? Two are addressed here: the political system's response to noncompliance, and the political system's coercive resource potential.

How does the political system respond to noncompliance? The typical short-term response of the coercive apparatus is to increase the level of resources devoted to coercion: as noncompliance increases, so do coercive resources. When the police chief, or the F.B.I. director, or the head of the secret police is faced with higher levels of law breaking, a first response is to increase resources in an effort to enforce the law. Different political systems demonstrate different levels of sensitivity toward noncompliance, and sensitivity varies across different laws. Thus the marginal effect of noncompliance upon the coercive response of the law enforcement agency varies in magnitude but it is always positive. This effect upon the law enforcement agency may be written as: dN_t, where d characterizes the level of sensitivity shown by a particular political system toward a particular crime.

The normal response of the law enforcement agency is, however, always mediated by the larger political system, and by the magnitude of noncompliance. Consider the case of marijuana laws: as marijuana law violations increase, efforts to enforce those laws increase as well, but at some point further coercive efforts become unfeasible. And thus, at some times and in some places, the enforcement effort is overwhelmed and aborted: a vivid example being rock concerts after the mid 1960s.

Interesting evidence along these lines comes from the work of Kuklinski and Stanga (1979). In a convincing analysis of California superior courts they show that sentencing severity in county marijuana cases changed as a function of the county vote in a state-wide marijuana initative. Counties that voted more heavily to remove criminal penalties for the personal use of marijuana were likely to realize reduced sentencing severity on the part of superior court judges. From the available evidence it is impossible to determine whether this relationship is due to (1) the potential for noncompliance, or (2) reduced levels of legitimacy. The important point is that the political system regularly intervenes to retreat from the enforcement of its own laws as the direct result of public behavior that undermines the enforcement effort.

Once again, it is important to make clear the point that such a retreat is not only relevant to laws that are peripheral to the survival of political systems. Even repressive regimes come up against limits in their ability to enforce unpopular laws, even when the laws are central to the continuation of the regime. (Witness the recent examples of South Korea, and the Phillipines under Marcos.) As a limiting case, even the most repressive regime can no longer enforce repressive laws when no one is willing to operate the instruments of repression.

In summary, mass noncompliance is capable of producing less rather than more coerciveness on the part of the political system, and thus it is reasonable to define a noncompliance saturation factor. At levels of noncompliance below the saturation point, M, higher levels of noncompliance produce more coercion, albeit at a decreasing rate. At levels of noncompliance above the saturation point, higher levels of noncompliance produce lower levels of coercion at an increasing rate. Thus M measures a system's insulation from noncompliance: Low levels for M indicate a political system that is easily saturated by noncompliance. High levels for M indicate a political system that is not easily dissuaded from its effort to enforce compliance -- a system that can withstand high levels of noncompliance without being saturated and forced to abort its enforcement effort.

This is not to suggest that mass noncompliance always leads to a decrease in coerciveness -- one has only to recall the practice of the Pinochet regime in imprisoning thousands of its opponents in a soccer stadium. At the same time, however, even the most repressive regimes are likely to have at least theoretical limits to their coercive potential. There is a point of noncompliance beyond which the state must retreat from its efforts to enforce its own laws, and thus it must constrain the efforts of the law enforcement agencies. As a practical matter, for many regimes, these limits may lie beyond the realm of the probable. In terms of the model being constructed here, M might be very large.

These observations may be combined into the following symbolic representation of the general effect of noncompliance upon the rate at which coercive resources change: $dN_t(M-N_t)$, where M characterizes the level of noncompliance that saturates the capabilities of the system, and d denotes the sensitivity of the political system to noncompliance. Thus, the general effect of noncompliance is specified as being the interdependent product of

(1) the normal response of law enforcement agencies to noncompliance, and (2) the political system's noncompliance saturation point.

The parameter M is measured in the same metric as N_t: As N_t approaches M from below, the effect on coercion goes from positive to zero. As N_t becomes increasingly larger than M, the effect on coercion becomes increasingly negative. The size of M is crucial: Democratic regimes and laws that are peripheral to the survival of political systems lead to small values for M, while totalitarian regimes and laws crucial to a system's survival lead to large values. If all else is equal, democratic systems should be overwhelmed by lower levels of noncompliance. Breaking a law is, in a sense, another form of political expression and democracies are by definition more responsive to popular control. Similarly, if all else is equal, noncompliant behavior which threatens the survival of a regime will produce larger values for M. Thus, just as with d, M is crime or offense specific.

The coercive response of the political system is not only constrained by mass behavior, it is also constrained by the availability of resources. No political system has unlimited resources, and the coercive function of the political system must compete with other functions for the resources that are available. Borrowing from the vocabulary of the budgeting literature, political systems develop conceptions of "fair shares" in the distribution of resources (Wildavsky, 1974). All else being equal, an agency or function that receives more than its "fair share" will have its allocation lowered, and one that receives less will have its allocation raised.

The level of attentiveness to resource shares is likely to vary across political systems and functions. A political system that is fiscally strained must pay great heed to fair shares because it cannot afford to do otherwise. At the same time, a system that is very aggressive in its attitudes toward law enforcement will be less likely to pay much attention to fair shares when the coercive function is being considered.

These observations are summarized as a resource strain factor, and the effect of resource strain upon the rate at which coercive resources change may be represented as: $f(S-P_t)$, where S characterizes the equilibrium level of resources devoted to coercion in the absence of noncompliance, and f indexes the resource sensitivity of the political system as it affects law enforcement.

The status of S requires extended attention. Because S is the equilibrium level of coercive resources absent noncompliance, it is directly interpreted as the coercive function's fair share of political system resources, determined internally by the political system. That is, S represents a fundamental, enduring, before the fact committment to the coercive function on the part of the political system. And thus S is accurately defined as the system's investment in preemptive coercion. For some laws S will be very high, and for others it will be very low. A system's investment in preemptive coercion also depends upon the nature of the system: open societies spend less on preemptive coercion than closed societies.

The importance of S as a practical matter is largely a function of f. Larger values of f -- higher levels of resource sensitivity or resource strain -- make it (1) more difficult for P_t to exceed S, and (2) lead to faster returns toward S in the event that P_t is pushed above S by higher levels of noncompliance. (See Cortes Przeworski, and Sprague (1974) and Huckfeldt (1983) for discussions of response time and system memory.)

This description of the coercive logic may be summarized in three statements. First, the normal short-term response by law enforcement agencies to higher levels of noncompliance is an increased effort to enforce the law.

Second, this normal response is mediated by the larger political system, so that increases in noncompliance beyond some saturation point force the political system to curtail its enforcement effort. Third, political systems establish basic, long-term committments of resources to the coercive function, and they make some attempt to keep short-term expenditures in line with those allocations.

These statements are translated into a mathematical form as: $dP_t/dt = R_tP_t$, where dP_t/dt is the instantaneous change in coercive resources relative to time and R_t denotes the time dependent rate operating upon coercive resources. This rate may be, in turn, expressed as an additive function of resource strain and the short-term response to noncompliance. Thus,

$$dP_t/dt = [f(S-P_t) + dN_t(M-N_t)]P_t. \qquad (2)$$

The logic of the dynamic interdependence between coercion and compliance is contained in Eqs. 1 and 2. The mathematical form of this argument lies in a social science tradition that traces to Richardson's (1960) analysis of arms races (also see: Gillespie, Zinnes, Tahim, Schrodt, and Rubison, 1977), and a life science tradition that traces to the early work of Lotka and Volterra and the more recent and very useful work of Rosenzweig and MacArthur (1963), Rosenzweig (1969), Maynard Smith (1974), May (1974) and Gilpin (1975). Because the model's systems states are coupled in their movement through time, the model must be analyzed in a manner taking account of this simultaneity.

The Analytic Framework

A framework is established in this section for the analysis of the model. First, system equilibria are defined for the relationship between coercion and noncompliance. Second, a method is outlined for determining dynamics and stability in the relationship between coercion and noncompliance.

A central conceptual device for a dynamic system such as that portrayed in Eqs. 1 and 2 is equilibrium: the point at which change is absent, both in terms of coercion and in terms of noncompliance. A systematic consideration of various equilibria is furthered by restating the model as:

$$dN_t/dt = [a(L-N_t) - bP_t]N_t \quad (3)$$

$$dP_t/dt = [f(S-P_t) + dMN_t - dN_t^2]P_t. \quad (4)$$

First notice that if P_t is set to zero, then dP_t/dt equals zero as well. What happens to N_t? It converges toward L, and thus a first equilibrium is seen to be: $E_1=(L,0)$, the point at which coercion is absent and anyone who might disobey the law does indeed disobey. Similarly, a second equilibrium is easily seen to be: $E_2=(0,S)$, the point at which noncompliance is absent and coercion tracks to S. Finally, note that a third equilibrium is: $E_3=(0,0)$. This is perhaps the formal statement for Eden -- no one has thought to break the law, and no one else has considered the need to enforce it.

All three equilibria are degenerative in the sense that one or both system states reach zero. E_1 through E_3 signify (1) runaway noncompliance in the absence of coercion, (2) the preemptive level of coercion in the absence of noncompliance, and (3) Eden -- full compliance without coercion. While all of these qualify as equilibria, all might also be unstable. In terms of the first equilibrium (E_1) it is highly unlikely that a political system would continue to tolerate noncompliance with important laws, unless the political system had been overwhelmed by noncompliance and forced to abort its enforcement effort. Indeed, if this equilibrium <u>is</u> stable, it signifies that a political system's effort to enforce its law has failed.

Many political systems aspire toward alternatives that lie between these three degenerative equilibria, where coercion and noncompliance are balanced in a manner that may be either durable or tenuous. These alternative equilibria are located by setting Eqs. 3 and 4 to zero ($dP_t/dt = dN_t/dt = 0$), and by rearranging the result to obtain:

$$P_t = (a/b)L - (a/b)N^* \quad (5)$$

$$P^* = S + (d/f)MN_t - (d/f)N_t^2. \quad (6)$$

Eq. 5 defines a straight line which yields equilibrium in N_t -- symbolized as N^*. Eq. 6 defines a quadratic which yields equilibrium in P_t -- defined as P^*. Both equations are defined in the same plane, and any point of intersection satisfies the equilibrium conditions for both N^* and P^*. That is, a point of intersection yields a system equilibrium.

In order to maintain the interpretation of the model, all parameters must be positively valued and this in turn maintains system motion in the first (positively valued) quadrant of the N,P plane -- the only quadrant that provides meaningful values for the system states. This is no assurance, however, that an intersection actually occurs within the first quadrant. Indeed, the scenarios which follow show that zero, one, or two such intersections may occur, and these intersections may either produce (locally) stable equilibria, or thresholds which demarcate stability from instability.

The implication is clear: if an intersection does not occur then a nonzero equilibrium is nonexistent and either noncompliance or coercion will be driven from the system. Alternatively, if an intersection does occur, it may or may not provide a stable balance between coercion and noncompliance.

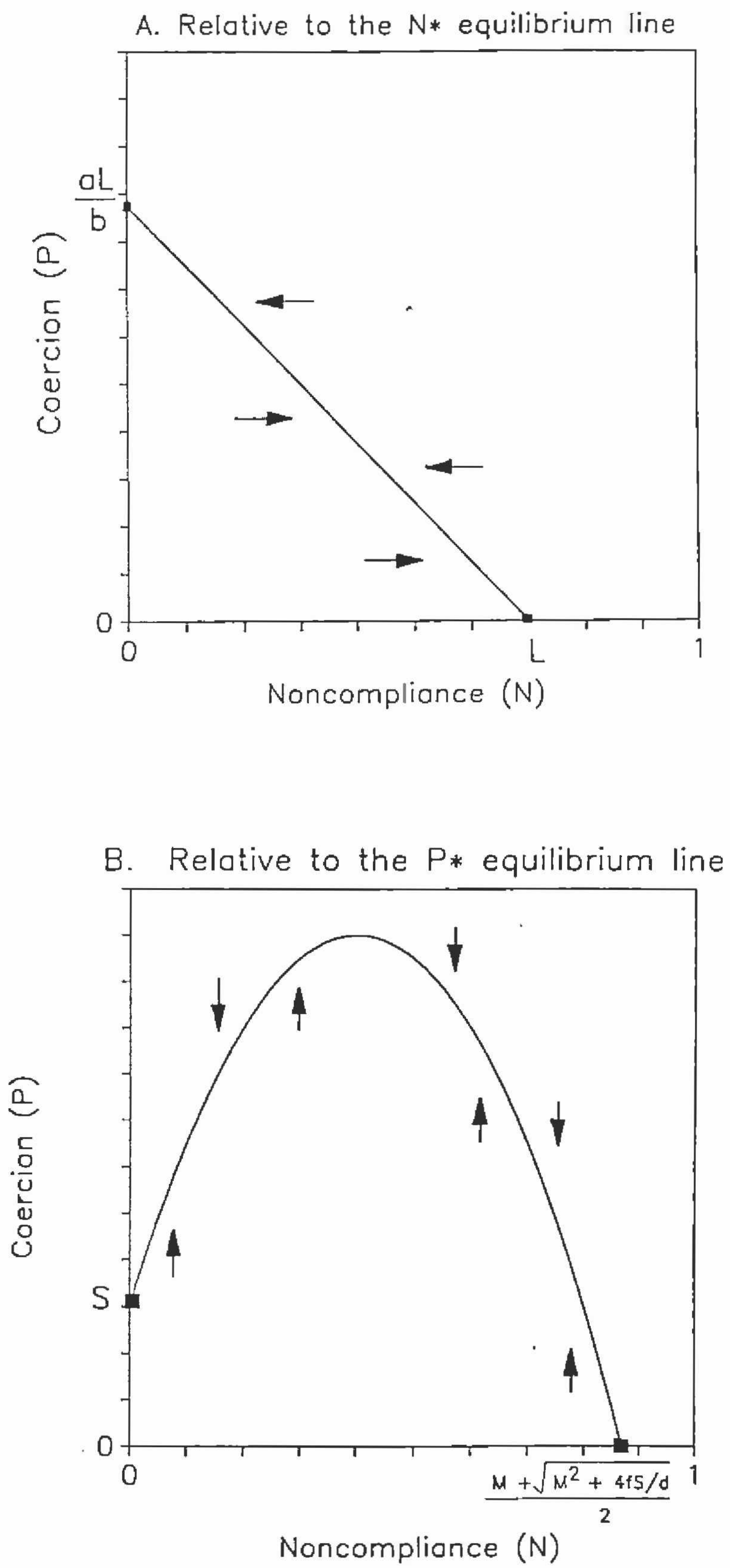
Thus it is important to develop a technology that allows a determination of system motion in the N,P plane (Maynard Smith, 1968; Huckfeldt, Kohfeld, and Likens, 1982). Figure 1a graphs the N^* equilibrium line of Eq. 5, and Figure 1b graphs the P* equilibrium line of Eq. 6. The trajectories shown in Figure 1a illustrate that first quadrant movement in N_t is always in the direction of the N* equilibrium line, and Figure 1b illustrates that first quadrant movement in P_t is always in the direction of the P* equilibrium line. Thus, simultaneous motion in P and N can be determined by taking account of both the vertical and horizontal "pulls" of each equilibrium line. In this manner the process can be allowed to unfold in time, where time is represented as a path in the N,P plane. (See the appendix for a more systematic development.)

Several other features of Figure 1 require attention because they serve as the basis for the analysis that follows. The N* equilibrium line intersects the P axis at: (a/b)L, and it intersects the N axis at: L. The P* equilibrium line intersects the P axis at: S. It intersects the N axis at two points, but only one point lies above zero: $[M + \sqrt{M^2 + 4fS/d}\,]/2$. Finally, the P* equilibrium line reaches its maximum in P when: N=M/2. These features of the equilibrium lines allow the development of several different scenarios, and the derivation of several results regarding the dynamics of coercion and noncompliance. The first scenario is concerned with the potential success of preemptive strategies in efforts to control noncompliance.

The Potential for Preemptive Success

This analysis begins by considering two scenarios, both of which might result in a successful preemptive strategy for eliminating noncompliance. In one scenario the preemptive strategy always succeeds but in the other it might fail, for reasons that are not entirely obvious. These two scenarios are shown in Figure 2. In Part A of the Figure there is no intersection between the two equilibrium lines. Thus a nonzero equilibrium is

Figure 1. Movement in the N,P plane.



not present, and motion in the system is toward: $E_1=(0,S)$. That is, noncompliance is driven out of the system, and coercion is maintained at its preemptive, fair share equilibrium. An intersection is present in Part B of the figure, but it is unstable, producing a threshold effect. Levels of noncompliance and coercion within a particular subarea of the plane produce runaway noncompliance and the demise of the coercive effort. Levels of noncompliance which lie outside this area produce movement toward E_1, and the elimination of noncompliance. The alternatives in this second scenario are either the complete elimination of noncompliance or runaway noncompliance, and only slight differences in initial conditions or minor exogenous shocks to the system might determine the result.

What do the two scenarios share in common? In both instances:

$$S > (a/b)L. \qquad (7)$$

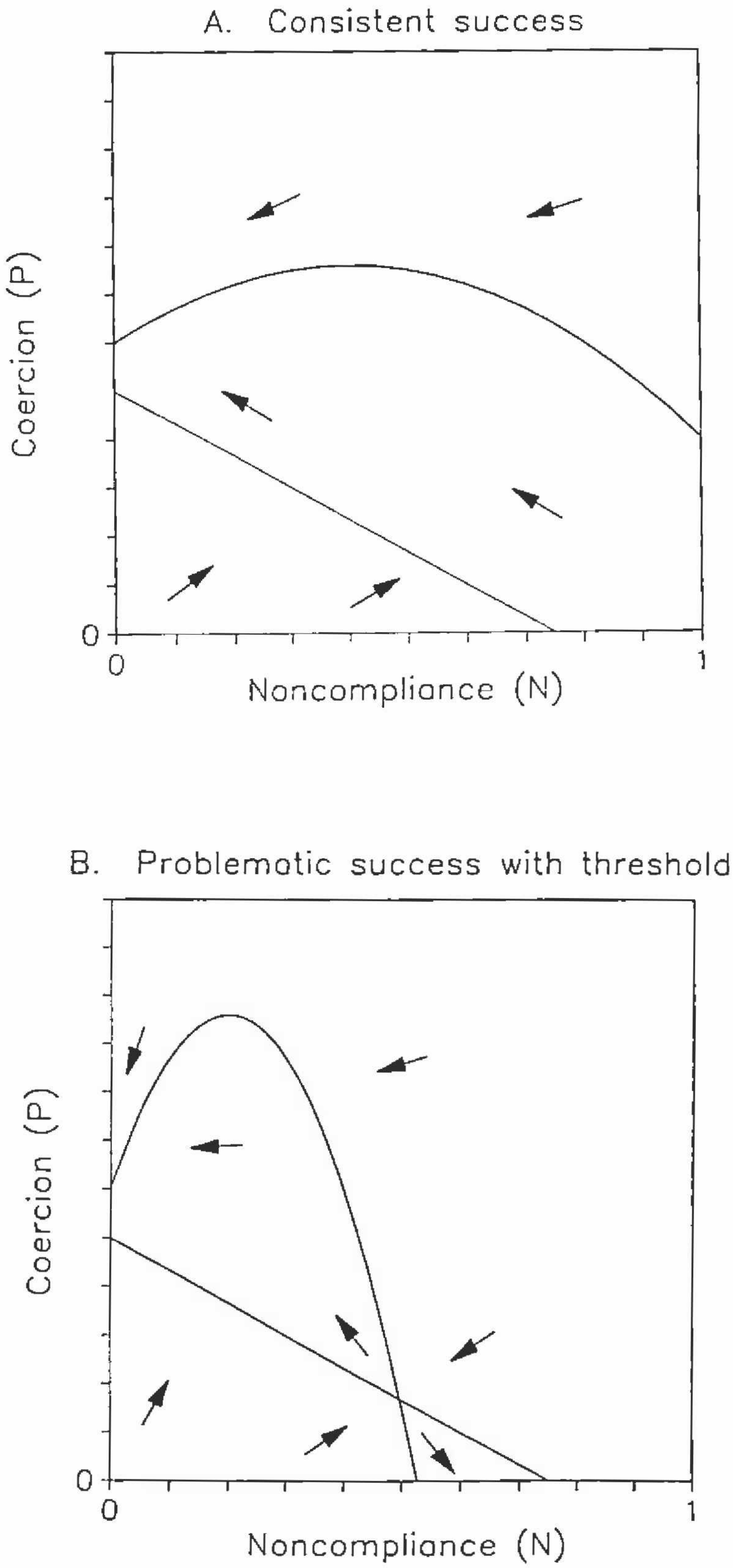
And it can be shown that this condition must hold if noncompliance is to be completely eliminated. This means that a political system is more likely to drive out noncompliance by (1) investing heavily in preemptive coercion (S), and (2) increasing the effectiveness of its enforcement activity (b). The potential for eliminating noncompliance is also furthered by: (1) a smaller proportion of citizens who are potentially willing to break the law (L), and (2) a lower level of profit (a) from noncompliance.

What is different about the two scenarios? In the first,

$$[M + \sqrt{M^2 + 4fS/d}\,]/2 > L. \qquad (8)$$

In the second scenario this inequality is reversed. First notice that Ineq. 8 <u>must</u> hold if M is greater than L -- if the noncompliance saturation point lies beyond the potential level of noncompliance. Thus, as legitimacy increases (as L decreases), and as the ability to withstand noncompliance increases, so does the potential for eliminating noncompliance. Indeed, a guarantee of

Figure 2. Preemptive strategies.



total compliance can be written wholly in terms of: preemptive coercion, coercion effectiveness, noncompliance profitability, system legitimacy, and the system's noncompliance saturation point. This guarantee leaves out any consideration of the political system's short-term response to noncompliance (f and d).

Can the preemptive strategy be guaranteed to succeed (can Ineq. 8 be satisfied) if potential noncompliance (L) exceeds the noncompliance saturation point (M)? Possibly, but herein lies an irony: the potential for <u>consistently</u> eliminating noncompliance is furthered by exercising <u>restraint</u> in the short-term response to noncompliance! Notice that Ineq. 8 is more likely to be satisfied if f grows in size and d declines in size. But this means that the preemptive strategy is more likely to succeed if the political system is unwilling or unable to increase its level of coercive resources above its fair share level, and if the system is less responsive to noncompliance in the short-term. (For interesting and complementary results generated using a different model and different assumptions see Salert and Sprague, 1980.)

Why? The exercise of coercion is risky business because it has inherent built-in limitations. Political systems that are willing and able to respond aggressively to noncompliance run the inherent risk of making themselves more vulnerable to the noncompliance saturation point. In graphic terms, the P^* line declines more rapidly after reaching its maximum, and thus it is more likely to intersect with the N^* line, producing a threshold.

In summary, the preemptive strategy is more likely to succeed in eradicating noncompliance if the political system invests heavily in a sustained fair share for the coercive apparatus, but then exercises restraint in its short-term response to noncompliance. Not only can such a political system <u>afford</u> patience, but patience pays its own dividends.

Controlling Noncompliance without Preemption

A successful preemptive strategy imposes potentially significant costs, especially in political systems with low levels of legitimacy (high L), and where the profitability of noncompliance (a) is high. First, the burden of a large resource share for the coercive apparatus must be maintained. Second, and perhaps more important, many people do not find it congenial to live in a police state, even if the coercive apparatus does exercise restraint in its response. The question then arises: what are the prospects for controlling noncompliance, absent a preemptive strategy?

The absence of a preemptive strategy may be represented formally by reversing Ineq. 7 so that:

$$(a/b)L > S. \qquad (9)$$

In words, the preemptive strategy is abortive if the fair share for coercion is less than potential noncompliance weighted by the ratio of noncompliant profitability to coercion effectiveness. Statements such as these are, of course, relative. It is possible to pursue a preemptive strategy with only a small level of S, so long as L is also very small. Conversely, it may not be possible to pursue the preemptive strategy even if S is very large, depending upon the magnitude of L.

The absence of a preemptive strategy has at least two consequences that not a few law abiding citizens might find congenial: (1) lower taxes, or more resources for other functions of the political system, and (2) the potential for a more open society. Is such a society doomed to rampant lawlessness?

Two scenarios, both of which either forsake or fail to attain the preemptive strategy, are shown in Figure 3. In Part A of the figure a stable equilibrium is attained between coercion and noncompliance. Any displacement away from this equilibrium is self-extinguishing. While noncompliance is not driven out of the system, it is contained at a level somewhere between complete compliance ($N^*=0$), and the maximum possible level of noncompliance ($N^*=L$).

In contrast, Part B of Figure 3 shows runaway noncompliance and the demise of the coercive effort. Noncompliance reaches its maximum level and coercive resources go to zero. At best (or worst) this signifies that the political system has given up enforcing an unpopular law. At worst (or best) it signifies the complete unravelling of the political system.

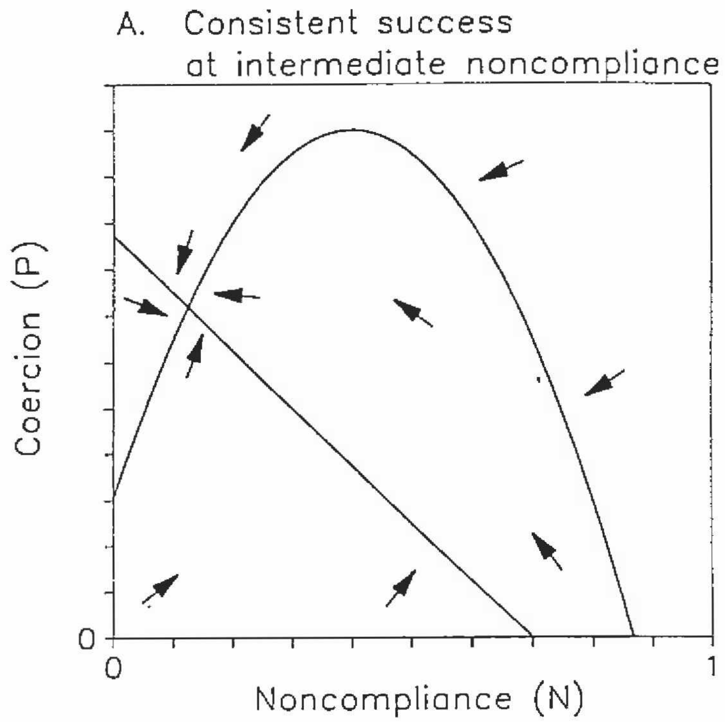
Ineq. 9 holds for both scenarios. How are the scenarios different? In the first instance of the stable equilibrium, where noncompliance is successfully controlled, it is the case that:

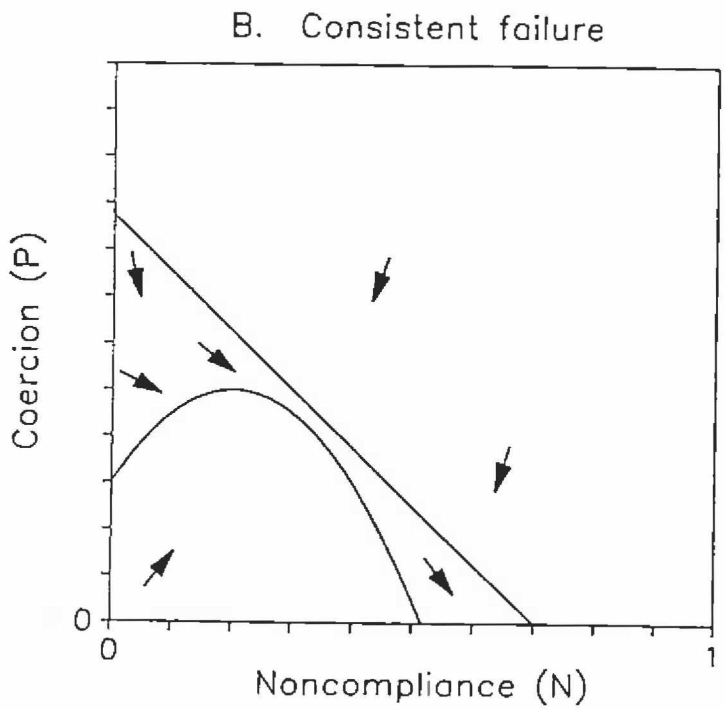
$$[M + \sqrt{M^2 + 4fS/d}\,]/2 > L. \qquad (10)$$

A reversal of Ineq. 10 will not necessarily produce the pattern of runaway compliance shown in Figure 3b; an alternative outcome, shown below in Figure 4, might also be obtained. For present purposes it is sufficient to observe that Ineq. 10 guarantees a stable equilibrium.

Recall that Ineq. 10 is the same condition that separated the consistently successful preemptive strategy from the preemptive strategy where success was both problematical and fragile. As before, success is guaranteed if $M > L$: if the point of noncompliance saturation is greater than the highest possible level of noncompliance, thereby making saturation impossible. And as before, restraint in the short-term response to noncompliance is the best policy. Even if $L > M$, restraint in the short run may be sufficient to generate stability. In mathematical terms, such restraint may be sufficient to satisfy Ineq. 10.

Figure 3. Non-preemptive strategies.





The important point is as follows: if a system's point of noncompliance saturation (M) exceeds the highest possible level of noncompliance that might be experienced (L), then either a preemptive or a non-preemptive strategy will be successful. The preemptive strategy will drive noncompliance out of the system, and the non-preemptive strategy will control noncompliance at some intermediate level. If this inequality is reversed, if L exceeds M, then the political system is generally well advised to exercise great restraint in its response to noncompliance. The dynamic logic of both scenarios suggests that runaway noncompliance might be avoided if the short-term response of the political system is restrained. An aggressive response only serves to make the political system more vulnerable to the noncompliance saturation point.

Desperate Strategies for Desperate Circumstances

Does it ever make good sense for the political system to react aggressively to noncompliance? To continue with the scenario of Figure 3b, suppose that the potential for noncompliance (L) is so much larger than the noncompliance saturation point (M) that no amount of patience and restraint can remedy the situation and bring noncompliance under control. What is a police chief to do?

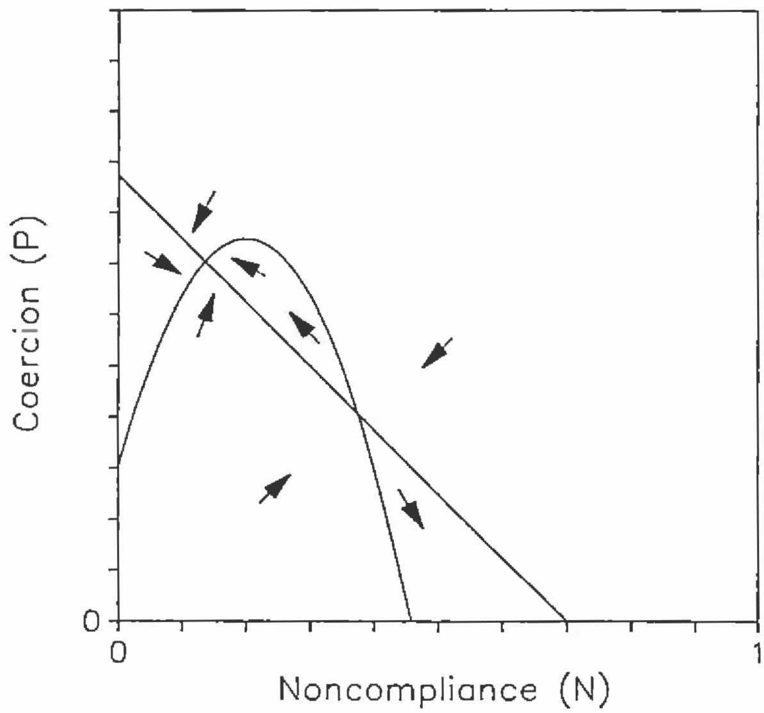
First notice that this situation ($M \ll L$) suggests a political system with profound difficulties. It means that either (1) legitimacy is very low, or that (2) the system is very vulnerable to even minor levels of noncompliance. Fortunately or unfortunately, such situations are not rare events because vulnerability and legitimacy are likely to be inversely correlated. As citizens become increasingly willing to break the law, they also become less willing to see the political system enforce it. Thus, as L increases, M decreases, and vice versa.

These considerations lead to a final scenario, depicted in Figure 4. The two inequalities that generated Figure 3b hold for this figure as well. In graphic terms, the difference between the two figures lies in the height of the hump for the P^* equilibrium line. In Figure 3b its maximum lies below the N^* equilibrium line, but in Figure 4 its maximum lies above the N^* equilibrium line, and thus two equilibria are produced.

As Figure 4 shows, the left hand side equilibrium is locally stable, so that displacements in its neighborhood produce movement back toward equilibrium. The right hand side equilibrium is inherently unstable: any displacement away from this equilibrium leads either to the left hand side equilibrium or to the demise of coercion and runaway noncompliance. Thus, the second equilibrium produces a threshold that lies between successful control of noncompliance on the one hand and system degeneration on the other. In this case, as in the threshold phenomenon of Figure 2b, initial conditions become profoundly important. For example, if a new law is passed, initial levels of compliance may very well determine whether the law can be successfully enforced, or whether runaway noncompliance is the inevitable result (Likens and Kohfeld, 1983). Similarly, initial levels of enforcement activity may very well determine whether the enforcement effort succeeds or fails.

Furthermore, these models are deterministic, but reality includes a stochastic component that produces random variation in time paths. Thus, a deterministic trajectory that is headed toward a stable equilibrium may be pushed over the threshold by a random event, thereby leading to runaway noncompliance and failure on the part of the enforcement effort. In summary,

Figure 4. Non-preemptive strategy: Problematic success with threshold.



while the political system of Figure 4 is better situated than the political system of Figure 3b, the advantage may not be enormous, and the degenerative outcome may be the same. (For several interesting discussions of threshold phenomena see Schelling 1978.)

What separates Figure 4 from Figure 3b? A double equilibrium must be produced if the maximum value of P for the P* equilibrium line lies above the N* equilibrium line along a line drawn perpendicular to the N axis. This will be the case if:

$$S + dM^2/4f > a(2L-M)/2b \qquad (11)$$

The left hand side of Ineq. 11 is the maximum value of P obtained by the P* equilibrium line, where the maximum in P is obtained at: N=M/2. The right hand side of Ineq. 11 is the value of P obtained by the N* equilibrium line at the same value of N -- M/2. Ineq. 11 is <u>not</u> the general condition for the existence of intersections between the two equilibrium lines. It is possible for intersections to occur even when Ineq. 11 is not satisfied. However, as L becomes increasingly larger than M -- as circumstances become more desperate for a political system -- it becomes increasingly unlikely that intersections will occur unless Ineq. 11 is satisfied.

Several features of Ineq. 11 merit attention. First, notice that this condition reverses many of the conditions that produced the present predicament. For example, higher levels of preemptive coercion help to satisfy the inequality, but a low level of preemptive coercion relative to system legitimacy, is largely responsible for producing this desperate scenario. All of which is to indicate that, by definition, this condition is only likely to be satisfied in a political system where legitimacy is <u>very</u> low, and levels of preemptive coercion are correspondingly high -- a political system that is both desperate <u>and</u> repressive. Second, notice that Ineq. 11 is more likely to be satisfied by increasing d and reducing f: by responding aggressively to noncompliance in the short-term, and by allowing flexibility in the fair share for coercive resources.

In short, the last hope for desperate leaders in desperate circumstances is to respond in an aggressive fashion to noncompliance, both in the long-term and in the short-term. But the benefits of such a strategy must be seen within the desperate context that made it beneficial, and the tenuous situation that is produced even if the strategy is successful.

Conclusion

This effort has drawn a distinction between long-term and short-term responses to noncompliance. A political system's level of preemptive coercion -- coercive resources in the absence of noncompliance -- constitutes a long-term response, either to noncompliance, or to the potential for noncompliance. It represents the enduring commitment that a political system makes, independent of particular noncompliance levels, to coercion as a device for securing compliance. The resulting magnitude of the coercive apparatus must be acknowledged as a central feature of any society. In this sense it represents a central institutional ingredient within political systems.

This long-term response is usefully contrasted to a short-term response: the actual level of resources committed to enforcement activities in response to changing levels of noncompliance. Short-term responsiveness depends both upon the fiscal capacities of the political system, and upon system aggressiveness in responding to noncompliance. The short-term response is likely to be most pronounced for systems that are: (1) willing and able to exercise flexibility in adjusting coercive resources, and (2) aggressive in response to increased levels of noncompliance.

The analysis undertaken here produces a nonobvious result: short-term restraint coupled with long-term aggressiveness offers the best potential for eliminating noncompliance. That is, only a preemptive strategy can succeed in eradicating noncompliance, and such a strategy is furthered by exercising short-term restraint in response to noncompliance. Short-term aggressiveness has the unintended consequence of making the political system more vulnerable to its point of noncompliance saturation, and hence more likely to be overwhelmed by noncompliance.

Furthermore, even in situations where a preemptive strategy is not pursued, short-term restraint will generally afford the best strategy for controlling noncompliance at some intermediate level. Only under the most desperate circumstances, in situations where a political system is both vulnerable to noncompliance and suffering from a low level of legitimacy, do increased levels of short-term aggressiveness pay any dividends. And even in these situations, the success of the strategy is problematic.

Finally, under some circumstances short-term restraint is guaranteed to result in the successful control of noncompliance. In political systems distinguished by high levels of legitimacy and by an ability to withstand high levels of noncompliance, short-term restraint is certain to keep noncompliance within manageable bounds, regardless of whether a system pursues a preemptive or nonpreemptive long-term strategy.

In summary, this analysis has important consequences for democratic prospects. Not only does it suggest that short-term coercive restraint is likely to pay practical dividends, but it also suggests that noncompliance with unpopular laws might be controlled without a fundamental institutional committment to excessive levels of preemptive coercion. The success of such restraint depends, inevitably, upon the system's level of legitimacy. But increasing legitimacy is not an impossible task for democratic systems. Indeed, it should be one of the things they do best.

Appendix

This appendix gives more detailed attention to Figure 1, the graphical framework for determining motion in the N,P plane. Each equilibrium line demonstrates the set of points at which a system state, considered separately, must be at rest. Taken together, any intersection provides a point where both states must be at rest.

Define two functions R_p and R_n, and recall that coercion is at equilibrium when:

$$R_p = gS - gP + dMN - dN^2 = 0,$$

and noncompliance is at equilibrium is at equilibrium when:

$$R_n = aL - aN - bP = 0.$$

These functions do not only determine equilibria, however, they also determine the direction of change in the respective system states. When R_p or R_n is less than zero, a decrease is produced in the corresponding system state. When R_p or R_n is greater than zero an increase is produced in the corresponding system state.

In the discussion that follows, each equilibrium line is at first considered separately to determine the effect of deviations from the equilibrium line upon one dimensional movement in the phase plane. The area above and to the right of the noncompliance equilibrium line contains points that are higher on either P or N, thereby producing a negative R_n and a decrease in noncompliance. Correspondingly, the area below and to the left contains points that are lower on either P or N, thereby producing a positive R_n and an increase in noncompliance. Thus, horizontal movement in the N,P plane -- change in the level of noncompliance -- is attracted toward the noncompliance equilibrium line.

Calculating movement in the vertical direction -- change in noncompliance -- is more complex due to the nonlinearity of the coercion equilibrium line. As a first step consider change in R_p with respect to N:

$$\partial R_p/\partial N = dM - 2dN.$$

From this it follows that the effect of N is fundamentally transformed at: N=M/2. Increased levels of N below the maximum contribute toward positive change in P, and decreased levels contribute toward negative change in P. Conversely, increased levels of N above the maximum contribute toward negative change in P, and decreased levels contribute toward positive change. In contrast to the effect of N, increased values for P always contribute toward negative change in N, and decreased levels always contribute toward positive change.

As a second step it is useful to divide the P* equilibrium line vertically at its maximum: N=M/2. Now consider the area above the equilibrium line and to the left of the its maximum. This area contains points that are either higher on P or lower on N than the points on the equilibrium line, and this produces a decrease in P. The area above the equilibrium line and to the right of the maximum contains points that are either higher on P or higher on N than the points on the equilibrium line, and this also produces a decrease in P. The area below the equilibrium line and to the left of the maximum contains points that are either lower on P or higher on N, producing an increase in P. And finally the area below the equilibrium line and to the right of the maximum contains points that are either lower on P or lower on N, also producing an increase in P. Taken together, this all means that vertical motion above the coercion equilibrium line is downward, and vertical motion below the coercion equilibrium line is upward.

Any point on the plane is subject to a horizontal attraction and a vertical attraction. The strength of each attraction is a direct function of the distance from the respective equilibrium line. If a point is removed from both equilibrium lines, its instantaneous motion will generally lie in a diagonal direction. If a point lies at an intersection, it will not move at all. If a point lies on one equilibrium line but not on another, its instantaneous motion will either be at an angle of 0 or 90 degrees, depending upon which equilibrium line it is located on. This one-directional motion will only last for an instant, however, as the two-dimensional attraction is renewed as soon as the trajectory departs from the equilibrium line.

The analysis of this paper is not, in general, concerned with the nature of movement toward (or away from) equilibrium. In particular, the possibility of limit cycles is not explicitly considered, even though it would appear that the system considered here does not generate such motion. For discussions of these issues see May (1974).

References

Barry, Brian. M. 1970. Sociologists, Economists and Democracy. London: Collier-Macmillan.

Becker, Gary. S. 1968. Crime and Punishment: An Economic Approach, Journal of Political Economy, 76: 169-217.

Cortes, Fernando, Adam Przeworski and John Sprague. 1974. Systems Analysis for Social Scientists. New York: Wiley.

DeNardo, James. 1985. Power in Numbers: The Political Strategy of Protest and Rebellion. Princeton: Princeton University Press.

Gillespie, John V., Dina A. Zinnes, Philip A. Schrodt, G. S. Tahim, and R. Michael Rubison. 1977. An Optimal Control Model of Arms Races, American Political Science Review, 71: 226-244.

Gilpin, Michael E. 1975. Group Selection in Predator-Prey Communities. Princeton: Princeton University Press.

Granovetter, Mark. 1978. Threshold Models of Collective Behavior, American Journal of Sociology, 83: 1420-1443.

Gurr, Ted Robert. 1970. Why Men Rebel. Princeton: Princeton University Press.

Huckfeldt, R. Robert. 1983. The Social Context of Political Change: Durability, Volatility, and Social Influence, American Political Science Review, 77: 929-944.

Huckfeldt, R. Robert, C. W. Kohfeld, and Thomas W. Likens. 1982. Dynamic Modeling: An Introduction. Beverly Hills: Sage.

Kuklinski, James H. and John E. Stanga. 1979. Political Participation and Government Responsiveness: The Behavior of California Superior Courts, American Political Science Review, 73: 1090-1099.

Likens, Thomas W. and C. W. Kohfeld. 1983. Models of Mass Compliance: Contextual or Economic Approach? Political Methodology, 9: 377-406.

May, Robert M. 1974. Stability and Complexity in Model Ecosystems. Princeton: Princeton University Press.

Monas, Sidney. 1961. The Third Section: Police and Society in Russia under Nicholas I. Cambridge: Harvard University Press.

Olson, Mancur. 1971. The Logic of Collective Action: Public Goods and the Theory of Groups. Revised Edition. New York: Schocken Books.

Przeworski, Adam. 1974. Contextual Models of Political Behavior, Political Methodology, 1: 27-61.

Rosenzweig, Michael L. 1969. Why the Prey Curve has a Hump, American Naturalist, 103: 81-87.

Rosenzweig, M. L. and R. H. MacArthur. 1963. Graphical Representation and Stability Conditions of Predator-Prey Interaction, American Naturalist, 97: 209-223.

Richardson, Lewis F. 1960. Arms and Insecurity. Pittsburgh: Boxwood.

Salert, Barbara. 1976. Revolutions and Revolutionaries: Four Theories. New York: Elsevier.

Salert, Barbara and John Sprague. 1980. The Dynamics of Riots. Ann Arbor: Interuniversity Consortium for Political and Social Research.

Schelling, Thomas C. 1978. Micromotives and Macrobehavior. New York: Norton.

Smith, John Maynard. 1968. Mathematical Ideas in Biology. Cambridge: Cambridge University Press.

Smith, John Maynard. 1974. Models in Ecology. Cambridge: Cambridge University Press.

Tilly, Charles. 1978. From Mobilization to Revolution. Reading, Massachusetts: Addison-Wesley.

Wildavsky, Aaron. 1974. The Politics of the Budgetary Process. Second Edition. Boston: Little, Brown.

ZUMA-Arbeitsberichte

80/15 Gerhard Arminger, Willibald Nagl, Karl F. Schuessler
Methoden der Analyse zeitbezogener Daten. Vortragsskripten der ZUMA-Arbeitstagung vom 25.09. - 05.10.79

81/07 Erika Brückner, Hans-Peter Kirschner, Rolf Porst, Peter Prüfer, Peter Schmidt
Methodenbericht zum "ALLBUS 1980"

81/19 Manfred Küchler, Thomas P. Wilson, Don H. Zimmerman
Integration von qualitativen und quantitativen Forschungsansätzen

82/03 Gerhard Arminger, Horst Busse, Manfred Küchler
Verallgemeinerte Lineare Modelle in der empirischen Sozialforschung

82/08 Glenn R. Carroll
Dynamic analysis of discrete dependent variables: A didactic essay

82/09 Manfred Küchler
Zur Messung der Stabilität von Wählerpotentialen

82/10 Manfred Küchler
Zur Konstanz der Recallfrage

82/12 Rolf Porst
"ALLBUS 1982" - Systematische Variablenübersicht und erste Ansätze zu einer Kritik des Fragenprogramms

82/13 Peter Ph. Mohler
SAR - Simple AND Retrieval mit dem Siemens-EDT-Textmanipulations-programm

82/14 Cornelia Krauth
Vergleichsstudien zum "ALLBUS 1980"

82/21 Werner Hagstotz, Hans-Peter Kirschner, Rolf Porst, Peter Prüfer
Methodenbericht zum "ALLBUS 1982"

83/09 Bernd Wegener
Two approaches to the analysis of judgments of prestige: Interindividual differences and the general scale

83/11 Rolf Porst
Synopse der ALLBUS-Variablen. Die Systematik des ALLBUS-Fragen-programms und ihre inhaltliche Ausgestaltung im ALLBUS 1980 und ALLBUS 1982

84/01 Manfred Küchler, Peter Ph. Mohler
Qualshop (ZUMA-Arbeitstagung zum "Datenmanagement bei qualitativen Erhebungsverfahren") - Sammlung von Arbeitspapieren und -berichten, Teil I + II

84/02 Bernd Wegener
Gibt es Sozialprestige? Konstruktion und Validität der Magnitude-Prestige-Skala

84/03 Peter Prüfer, Margrit Rexroth
Erfahrungen mit einer Technik zur Bewertung von Interviewerverhalten

84/04 Frank Faulbaum
Ergebnisse der Methodenstudie zur internationalen Vergleichbarkeit von Einstellungsskalen in der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) 1982

84/05 Jürgen Hoffmeyer-Zlotnik
Wohnquartiersbeschreibung. Ein Instrument zur Bestimmung des sozialen Status von Zielhaushalten

84/07 Gabriele Hippler, Hans-Jürgen Hippler
Reducing Refusal Rates in the Case of Threatening Questions: The "Door-in-the-Face" Technique

85/01 Hartmut Esser
Befragtenverhalten als "rationales Handeln" - Zur Erklärung von Antwortverzerrungen in Interviews

85/03 Rolf Porst, Peter Prüfer, Michael Wiedenbeck, Klaus Zeifang
Methodenbericht zum "ALLBUS 1984"

86/01 Dagmar Krebs
Zur Konstruktion von Einstellungsskalen im interkulturellen Vergleich

86/02 Hartmut Esser
Können Befragte lügen? Zum Konzept des "wahren Wertes" im Rahmen der handlungstheoretischen Erklärung von Situationseinflüssen bei der Befragung

86/03 Bernd Wegener
Prestige and Status as Function of Unit Size

86/04 Frank Faulbaum
Very Soft Modeling: The Logical Specification and Analysis of Complex Process Explanations with Arbitrary Degrees of Underidentification and Variables of Arbitrary Aggregation and Measurement Levels

86/05 Peter Prüfer, Margrit Rexroth (Übersetzung: Dorothy Duncan)
On the Use of the Interaction Coding Technique

86/06 Hans-Peter Kirschner
Zur Kessler-Greenberg-Zerlegung der Varianz der Meßdifferenz zwischen zwei Meßzeitpunkten einer Panel-Befragung

86/07 Georg Erdmann
Ansätze zur Abbildung sozialer Systeme mittels nicht-linearer dynamischer Modelle

86/09 Heiner Ritter
Einige Ergebnisse von Vergleichstests zwischen den PC- und Mainframe-Versionen von SPSS und SAS

86/10 Hans-Peter Kirschner
Der Stichprobenplan zum Projekt ISSP 1985 und seine Realisierung

86/11 Günter Rothe
Bootstrap in generalisierten linearen Modellen

87/01 Klaus Zeifang
Die Test-Retest-Studie zum ALLBUS 1984 - Tabellenband

87/02 Klaus Zeifang
Die Test-Retest-Studie zum ALLBUS 1984 - Abschlußbericht

87/03 Michael Braun
ALLBUS-Bibliographie (6. Fassung, Stand: 30.06.87)

87/04 Barbara Erbslöh, Michael Wiedenbeck
Methodenbericht zum "ALLBUS 1986"

87/05 Norbert Schwarz, Julia Bienias
What Mediates the Impact of Response Alternatives on Behavioral Reports?

87/06 Norbert Schwarz, Fritz Strack, Gesine Müller, Brigitte Chassein
The Range of Response Alternatives May Determine the Meaning of the Question: Further Evidence on Informative Functions of Response Alternatives

87/07 Fritz Strack, Leonard L. Martin, Norbert Schwarz
The Context Paradox in Attitude Surveys: Assimilation or Contrast?

87/08 Gudmund R. Iversen
Introduction to Contextual Analysis

87/09 Seymour Sudman, Norbert Schwarz
Contributions of Cognitive Psychology to Data Collection in Marketing Research

87/10 Norbert Schwarz, Fritz Strack, Denis Hilton, Gabi Naderer
Base-Rates, Representativeness, and the Logic of Conversation

87/11 George F. Bishop, Hans-Jürgen Hippler, Norbert Schwarz, Fritz Strack
A Comparison of Response Effects in Self-Administered and Telephone Surveys

87/12 Norbert Schwarz
Stimmung als Information. Zum Einfluß von Stimmungen und Emotionen auf evaluative Urteile

88/01 Antje Nebel, Fritz Strack, Norbert Schwarz
Tests als Treatment: Wie die psychologische Messung ihren Gegenstand verändert

88/02 Gerd Bohner, Herbert Bless, Norbert Schwarz, Fritz Strack
What Triggers Causal Attributions? The Impact of Valence and Subjective Probability

88/03 Norbert Schwarz, Fritz Strack
The Survey Interview and the Logic of Conversation: Implications for Questionnaire Construction

88/04 Hans-Jürgen Hippler, Norbert Schwarz
"No Opinion"-Filters: A Cognitive Perspective

88/05 Norbert Schwarz, Fritz Strack
Evaluating One`s Life: A Judgment of Subjective Well-Being

88/06 Norbert Schwarz, Herbert Bless, Gerd Bohner, Uwe Harlacher, Margit Kellenbenz
Response Scales as Frames of Reference:
The Impact of Frequency Range on Diagnostic Judgments

88/07 Michael Braun
Allbus-Bibliographie
(7. Fassung, Stand: 30.6.88)

88/08 Günter Rothe
Ein Ansatz zur Konstruktion inferenzstatistisch verwertbarer Indices

88/09 Ute Hauck, Reiner Trometer
Methodenbericht
International Social Survey Program - ISSP 1987

88/10 Norbert Schwarz
Assessing frequency reports of mundane behaviors:
Contributions of cognitive psychology to questionnaire construction

88/11 Norbert Schwarz, B. Scheuring (sub.).
Judgments of relationship satisfaction: Inter- and intraindividual comparison strategies as a function of questionnaire strcuture

88/12 Rolf Porst, Michael Schneid
Ausfälle und Verweigerungen bei Panelbefragungen
- Ein Beispiel -

88/13 Cornelia Züll
SPSS-X. Anmerkungen zur Siemens BS2000 Version.

88/14 Michael Schneid
Datenerhebung am PC - Vergleich der Interviewprogramme "interv$^+$" und "THIS"

88/15 Norbert Schwarz, Bettina Scheuring
Die Vergleichsrichtung bestimmt das Ergebnis von Vergleichsprozessen:
Ist - Idealdiskrepanzen in der Partnerwahrnehmung

89/01 Norbert Schwarz, George F. Bishop, Hans-J. Hippler, Fritz Strack
Psychological Sources Of Response Effects in Self-Administered And Telephone Surveys

89/02 Michael Braun, Reiner Trometer, Michael Wiedenbeck,
Methodenbericht. Allgemeine Bevölkerungsumfrage der Sozialwissenschaften - ALLBUS 1988 -

89/03 Norbert Schwarz
Feelings as Information:
Informational and Motivational Functions of Affective States

89/04 Günter Rothe
Jackknife and Bootstrap:
Resampling-Verfahren zur Genauigkeitsschätzung von Parameterschätzungen

89/05 Herbert Bless, Gerd Bohner, Norbert Schwarz und Fritz Strack
Happy and Mindless?
Moods and the Processing of Persuasive Communications

89/06 Gerd Bohner, Norbert Schwarz und Stefan E. Hormuth
Die Stimmungs-Skala: Eine deutsche Version des "Mood Survey"
von Underwood und Froming

89/07 Ulrich Mueller
Evolutionary Fundamentals of Social Inequality, Dominance
and Cooperation

NONCOMPLIANCE AND THE LIMITS OF COERCION: THE PROBLEMATIC ENFORCEMENT OF UNPOPULAR LAWS*

Robert Huckfeldt
Indiana University
May 1988

ABSTRACT

This paper constructs a mathematical representation of the dynamic logic tying together coercion and noncompliance within political systems. Several concepts are central to the effort: system legitimacy, the long-term institutionalized investment of the political system in coercive resources, the vulnerability of political systems to mass noncompliance, and the short-term response of the political system to noncompliance. The analysis addresses a number of issues. What are the dynamic consequences of system legitimacy and of the political system's institutionalized investment in coercive resources? What are the consequences of its short-term coercive response to noncompliance? Under what circumstances will noncompliance be eliminated? Under what circumstances will it be controlled? Under what circumstances will political systems become overwhelmed by noncompliance? A central argument of this paper is that aggressive short-term responses to noncompliance are likely to be ineffective in controlling noncompliance, and they are likely to aggravate the vulnerability of political systems to mass noncompliance.

*This research has been supported through the generosity of Hartmut Esser, Peter Mohler and other colleagues at Zentrum für Umfragen, Methoden und Analysen in Mannheim. I am also grateful to a number of colleagues who have provided helpful suggestions along the way: Robert Muncaster, Vincent Ostrom, Barbara Salert, Peter Schmidt, John Sprague, Harvey Starr, Robert Stein, several anonymous reviewers, and a number of perceptive graduate students.

Note: This paper is forthcoming in Mathematical Modelling.

NONCOMPLIANCE AND THE LIMITS OF COERCION: THE PROBLEMATIC ENFORCEMENT OF UNPOPULAR LAWS

Robert Huckfeldt
Indiana University
May 1988

Some laws secure broad support and compliance within political systems because they are fundamental to the social order. Murder, extortion, theft, and kidnapping are broadly conceived as socially deviant behaviors, at least in nonrevolutionary circumstances, and the prohibition of these activities generates little controversy. In contrast, most political systems address an entirely different set of problems with respect to the enforcement of unpopular laws -- laws that proscribe activities not seen as deviant by large parts of the population. These laws are much more controversial because, in one way or another, the public disputes their validity. The enumeration of such laws is lengthy, and forbidden activities under various regimes range from (1) speeding on highways and smoking marijuana to (2) exercising free speech and participating in political assemblies. Unpopular laws are distinguished not by the seriousness of the activity they forbid, but rather by the level of public resistance they provoke. These laws generate an important political dimension because they involve the clashing interests of citizens and political systems. They produce a potential confrontation between the coercive power of the state and noncompliance on the part of the public.

This relationship between coercion and compliance is fundamental to politics, and to the viability of political systems. The maintenance of order and the survival of political systems require large scale compliance with politically established norms and legal mandates, and the state's coercive capacity is a central ingredient in securing compliance. Of course, citizens need not always be coerced, even to obey unpopular laws. Compliance is also secured voluntarily, absent even the threat of sanctions for noncompliance, and therein lies the central ingredient of system legitimacy. Furthermore, the coercive capacity of the state is limited, especially in democratic regimes, but also in political systems lacking popular control. At some level of noncompliance the coercive response of the state must reach its maximum and decline. In a democratic regime this is the very definition for popular control, and in a non-democratic regime it is the definition for revolution.

This paper represents the relationship between coercion and compliance in the form of a mathematical model. The model incorporates elements that are both dynamic and simultaneous in timing. The interdependence between coercion and compliance is viewed as a process that unfolds in time. Levels of coercion are preconditioned by previous levels of noncompliance, while levels of compliance are preconditioned by previous levels of coercion.

The mathematical representation is, in turn, employed to address several questions: Why do some efforts at enforcing unpopular laws fail, while others succeed? To what extent is success a function of system legitimacy and the response by citizens to sanctions for noncompliance? To what extent is success a function of the political system's long-term investment in coercive resources, and of its short-term coercive response to noncompliance? Under what circumstances can a political system obtain compliance with unpopular laws absent significant coercion? Finally, what are the strategic consequences of the limitations upon coercion as a strategy for securing compliance? How do these limitations affect the relationship between coercion and noncompliance, and the exercise of coercion by political systems?

A Characterization of Compliance

Why do people break laws? More importantly, why do some people quit obeying laws, and why do others quit breaking them? First, some people obey laws because they are laws, regardless of any penalties that might be levied against law breakers, and regardless of any benefits that might be obtained from noncompliance. Thus, there is some limit upon noncompliance under normal circumstances, and this limit is defined here as L: the proportion of citizens who might refuse to comply with a particular law (Likens and Kohfeld, 1983). Legitimacy is defined as the proportion who would never break the law under normal circumstances, 1-L.

Who then might become noncompliers? Let N_t denote the proportion of citizens who disobey the law at time t. It follows that the proportion of citizens eligible to become noncompliers at the next instant in time is equal to $L-N_t$, the proportion of citizens who might break the law, but who are currently obeying it. Thus L is defined as being fixed in time, while N_t varies through time. (In the dynamic representations of this paper, time subscripts are maintained on system states to emphasize the fact that they vary with time, and to call attention to the time invariant model parameters. All rates of change are instantaneous, however, and no time delays are built into the system.)

Compliers become noncompliers for two interdependent reasons: First, law abiding citizens are more willing to become lawbreakers to the extent that breaking a particular law is profitable, and the profitability of breaking a particular law is conceived as being constant in time (a). This is not to say that profitability is constant across crimes or across individuals. It is generally more profitable to break a speed law on a trip across Wyoming than it is to break a speed law on a trip to the grocery. Similarly it is more profitable for the employee who is late for work to break a speed law than it is for the employee who is on time. Thus the profitability of noncompliance (a) is defined as an average within a population relative to a particular law.

Profitability derives from numerous sources: some rioters steal color televisions from department stores in order to enjoy football games, while others enjoy the release of pent-up fury, while still others view their involvement as an instrumental act aimed at bringing down an oppressive political system. Indeed, economic analyses of compliance occupy a prominent place in the literature on crime and deterrence. A seminal work is that of Becker (1968), in which he argues that ". . . a useful theory of criminal behavior can dispense with special theories of anomie, psychological inadequacies, or inheritance of special traits and simply extend the economist's usual analysis of choice" (p. 170).

Profitability is not the only reason that law abiding citizens become law breakers, however. The individual level transformation from obeying to breaking the law is also more likely to occur if the general level of noncompliance is higher. Two different micro-level premises support this assertion. First, the rationally calculating, football-loving rioter might decide that it is safer to steal the television if many other people are stealing televisions, stereos and so forth. That is, the probability of individual apprehension decreases as the aggregate level of noncompliance increases (Granovetter, 1978). Thus, a social choice interpretation of individual behavior can be called upon to support the assertion that people are more likely to disobey a law to the extent that others also disobey the law.

Second, the importance of aggregate noncompliance for the individual decision to disobey a law might also be supported on the basis of a social structural interpretation of individual behavior. Some forms of noncompliance may be furthered through social interaction and social persuasion, and in these instances aggregate noncompliance is important as a source of social influence. At one extreme, revolutionary success requires revolutionary solidarity, and thus the revolution proceeds (diffuses) as revolutionaries are able to convert nonrevolutionaries. At some point, even many unwilling revolutionaries may be swept along by the social tide as the level of revolutionary activity increases. (For complementary analyses see Przeworski 1974, Tilly 1978, and DeNardo 1985: 45-46.)

Barry (1970) and Salert (1976) show that considerable difficulties occur when collective revolutionary activity is conceived in terms of individually based, maximizing behavior in the tradition of Olson (1971). These problems do not affect the present representation, however. The form of the model allows for either a social choice interpretation, or a social structure interpretation. One or the other will be more or less appropriate depending upon the form of noncompliance being considered. The important point is that both sets of micro-level premises support the importance of aggregate noncompliance as a factor affecting the individual's decision regarding compliance.

These two factors -- profitability and aggregate noncompliance -- are combined interactively as a product to form a rate at which compliers become noncompliers: aN_t. And thus the increase in noncompliance is defined as: $aN_t(L - N_t)$.

Why do law breakers become law abiding citizens? They respond to the coercive sanctions of the state, either because they are apprehended and punished, or because they fear such apprehension. These coercive sanctions may usefully be conceived as two separate components: the resources available to the coercive apparatus, and the effectiveness of the coercive apparatus. For example, Tsar Nicholas I invested heavily in coercive resources to silence the Decembrists and other political critics, but the effort was undermined by the ineffectiveness of the coercive apparatus. The effectiveness of the state's coercive effort is determined by a variety of factors: corruption within the system, institutional inefficiency, constitutional safeguards for citizens, and in the case of Russia's Third Section, the assignment of mutually contradictory and impossible tasks (Monas, 1961). These factors tend to be deeply ingrained within particular political systems, and thus effectiveness is treated as being constant in time with respect to a particular law (b).

In contrast, coercive resources are highly variable across time. A typical short-term response to higher levels of noncompliance is additional investment in coercive resources. Thus, coercive resources are defined relative to time as P_t, and the rate at which noncompliers become compliers is defined as an interaction between these two factors, written as the simple product: bP_t. This means, in turn, that the <u>decrease</u> in noncompliance is represented as: bP_tN_t.

As the works of Gurr (1970: chap. 8) and Salert and Sprague (1980) demonstrate, the public is not always encouraged to comply by increased levels of coercion. Increased coercion is capable of provoking outrage rather than restraint, thereby producing increased levels of noncompliance. This possibility generates important analytic consequences, but the focus of the present effort is upon variations in the response of the political system to noncompliance, rather than variations in the response of citizens to coercion.

Exploring both types of variations simultaneously lies beyond the bounds of this paper, and the present effort must be understood subject to the assumption that noncompliance is discouraged by increased coercion.

In summary, the net rate of <u>change</u> in noncompliance relative to time (dN_t/dt) is represented as the increase in noncompliance minus the decrease in noncompliance, or:

$$dN_t/dt = aN_t(L-N_t) - bP_tN_t. \qquad (1)$$

<u>A Characterization of Coercion</u>

What factors give rise to changes in the resource level devoted by the political system to coercion? Two are addressed here: the political system's response to noncompliance, and the political system's coercive resource potential.

How does the political system respond to noncompliance? The typical short-term response of the coercive apparatus is to increase the level of resources devoted to coercion: as noncompliance increases, so do coercive resources. When the police chief, or the F.B.I. director, or the head of the secret police is faced with higher levels of law breaking, a first response is to increase resources in an effort to enforce the law. Different political systems demonstrate different levels of sensitivity toward noncompliance, and sensitivity varies across different laws. Thus the marginal effect of noncompliance upon the coercive response of the law enforcement agency varies in magnitude but it is always positive. This effect upon the law enforcement agency may be written as dN_t, where d characterizes the level of sensitivity shown by a particular political ~~system~~ toward a particular crime.

The normal response of the law enforcement agency is, however, always mediated by the larger political system, and by the magnitude of

keine gute Notation, da $\frac{dN_t}{dt}$ in (1) andere Bedeutung, confusing!

noncompliance. Consider the case of marijuana laws: as marijuana law violations increase, efforts to enforce those laws increase as well, but at some point further coercive efforts become unfeasible. And thus, at some times and in some places, the enforcement effort is overwhelmed and aborted: a vivid example being rock concerts after the mid 1960s.

Interesting evidence along these lines comes from the work of Kuklinski and Stanga (1979). In a convincing analysis of California superior courts they show that sentencing severity in county marijuana cases changed as a function of the county vote in a state-wide marijuana initative. Counties that voted more heavily to remove criminal penalties for the personal use of marijuana were likely to realize reduced sentencing severity on the part of superior court judges. From the available evidence it is impossible to determine whether this relationship is due to (1) the potential for noncompliance, or (2) reduced levels of legitimacy. The important point is that the political system regularly intervenes to retreat from the enforcement of its own laws as the direct result of public behavior that undermines the enforcement effort.

Once again, it is important to make clear the point that such a retreat is not only relevant to laws that are peripheral to the survival of political systems. Even repressive regimes come up against limits in their ability to enforce unpopular laws, even when the laws are central to the continuation of the regime. (Witness the recent examples of South Korea, and the Phillipines under Marcos.) As a limiting case, even the most repressive regime can no longer enforce repressive laws when no one is willing to operate the instruments of repression.

In summary, mass noncompliance is capable of producing less rather than more coerciveness on the part of the political system, and thus it is reasonable to define a noncompliance saturation factor. At levels of noncompliance below the saturation point, M, higher levels of noncompliance produce more coercion, albeit at a decreasing rate. At levels of noncompliance above the saturation point, higher levels of noncompliance produce lower levels of coercion at an increasing rate. Thus M measures a system's insulation from noncompliance: Low levels for M indicate a political system that is easily saturated by noncompliance. High levels for M indicate a political system that is not easily dissuaded from its effort to enforce compliance -- a system that can withstand high levels of noncompliance without being saturated and forced to abort its enforcement effort.

This is not to suggest that mass noncompliance always leads to a decrease in coerciveness -- one has only to recall the practice of the Pinochet regime in imprisoning thousands of its opponents in a soccer stadium. At the same time, however, even the most repressive regimes are likely to have at least theoretical limits to their coercive potential. There is a point of noncompliance beyond which the state must retreat from its efforts to enforce its own laws, and thus it must constrain the efforts of the law enforcement agencies. As a practical matter, for many regimes, these limits may lie beyond the realm of the probable. In terms of the model being constructed here, M might be very large.

These observations may be combined into the following symbolic representation of the general effect of noncompliance upon the rate at which coercive resources change: $dN_t(M-N_t)$, where M characterizes the level of noncompliance that saturates the capabilities of the system, and d denotes the sensitivity of the political system to noncompliance. Thus, the general effect of noncompliance is specified as being the interdependent product of

(1) the normal response of law enforcement agencies to noncompliance, and (2) the political system's noncompliance saturation point.

The parameter M is measured in the same metric as N_t: As N_t approaches M from below, the effect on coercion goes from positive to zero. As N_t becomes increasingly larger than M, the effect on coercion becomes increasingly negative. The size of M is crucial: Democratic regimes and laws that are peripheral to the survival of political systems lead to small values for M, while totalitarian regimes and laws crucial to a system's survival lead to large values. If all else is equal, democratic systems should be overwhelmed by lower levels of noncompliance. Breaking a law is, in a sense, another form of political expression and democracies are by definition more responsive to popular control. Similarly, if all else is equal, noncompliant behavior which threatens the survival of a regime will produce larger values for M. Thus, just as with d, M is crime or offense specific.

The coercive response of the political system is not only constrained by mass behavior, it is also constrained by the availability of resources. No political system has unlimited resources, and the coercive function of the political system must compete with other functions for the resources that are available. Borrowing from the vocabulary of the budgeting literature, political systems develop conceptions of "fair shares" in the distribution of resources (Wildavsky, 1974). All else being equal, an agency or function that receives more than its "fair share" will have its allocation lowered, and one that receives less will have its allocation raised.

The level of attentiveness to resource shares is likely to vary across political systems and functions. A political system that is fiscally strained must pay great heed to fair shares because it cannot afford to do otherwise. At the same time, a system that is very aggressive in its attitudes toward law enforcement will be less likely to pay much attention to fair shares when the coercive function is being considered.

These observations are summarized as a resource strain factor, and the effect of resource strain upon the rate at which coercive resources change may be represented as: $f(S-P_t)$, where S characterizes the equilibrium level of resources devoted to coercion in the absence of noncompliance, and f indexes the resource sensitivity of the political system as it affects law enforcement.

The status of S requires extended attention. Because S is the equilibrium level of coercive resources absent noncompliance, it is directly interpreted as the coercive function's fair share of political system resources, determined internally by the political system. That is, S represents a fundamental, enduring, <u>before the fact</u> committment to the coercive function on the part of the political system. And thus S is accurately defined as the system's investment in <u>preemptive coercion</u>. For some laws S will be very high, and for others it will be very low. A system's investment in preemptive coercion also depends upon the nature of the system: open societies spend less on preemptive coercion than closed societies.

The importance of S as a practical matter is largely a function of f. Larger values of f -- higher levels of resource sensitivity or resource strain -- make it (1) more difficult for P_t to exceed S, and (2) lead to faster returns toward S in the event that P_t is pushed above S by higher levels of noncompliance. (See Cortes Przeworski, and Sprague (1974) and Huckfeldt (1983) for discussions of response time and system memory.)

This description of the coercive logic may be summarized in three statements. First, the normal short-term response by law enforcement agencies to higher levels of noncompliance is an increased effort to enforce the law.

Second, this normal response is mediated by the larger political system, so that increases in noncompliance beyond some saturation point force the political system to curtail its enforcement effort. Third, political systems establish basic, long-term committments of resources to the coercive function, and they make some attempt to keep short-term expenditures in line with those allocations.

These statements are translated into a mathematical form as: $dP_t/dt = R_tP_t$, where dP_t/dt is the instantaneous change in coercive resources relative to time and R_t denotes the time dependent rate operating upon coercive resources. This rate may be, in turn, expressed as an additive function of resource strain and the short-term response to noncompliance. Thus,

$$dP_t/dt = [f(S-P_t) + dN_t(M-N_t)]P_t. \quad (2)$$

The logic of the dynamic interdependence between coercion and compliance is contained in Eqs. 1 and 2. The mathematical form of this argument lies in a social science tradition that traces to Richardson's (1960) analysis of arms races (also see: Gillespie, Zinnes, Tahim, Schrodt, and Rubison, 1977), and a life science tradition that traces to the early work of Lotka and Volterra and the more recent and very useful work of Rosenzweig and MacArthur (1963), Rosenzweig (1969), Maynard Smith (1974), May (1974) and Gilpin (1975). Because the model's systems states are coupled in their movement through time, the model must be analyzed in a manner taking account of this simultaneity.

The Analytic Framework

A framework is established in this section for the analysis of the model. First, system equilibria are defined for the relationship between coercion and noncompliance. Second, a method is outlined for determining dynamics and stability in the relationship between coercion and noncompliance.

A central conceptual device for a dynamic system such as that portrayed in Eqs. 1 and 2 is equilibrium: the point at which change is absent, both in terms of coercion and in terms of noncompliance. A systematic consideration of various equilibria is furthered by restating the model as:

$$dN_t/dt = [a(L-N_t) - bP_t]N_t \quad (3)$$

$$dP_t/dt = [f(S-P_t) + dMN_t - dN_t^2]P_t. \quad (4)$$

First notice that if P_t is set to zero, then dP_t/dt equals zero as well. What happens to N_t? It converges toward L, and thus a first equilibrium is seen to be: $E_1=(L,0)$, the point at which coercion is absent and anyone who might disobey the law does indeed disobey. Similarly, a second equilibrium is easily seen to be: $E_2=(0,S)$, the point at which noncompliance is absent and coercion tracks to S. Finally, note that a third equilibrium is: $E_3=(0,0)$. This is perhaps the formal statement for Eden -- no one has thought to break the law, and no one else has considered the need to enforce it.

All three equilibria are degenerative in the sense that one or both system states reach zero. E_1 through E_3 signify (1) runaway noncompliance in the absence of coercion, (2) the preemptive level of coercion in the absence of noncompliance, and (3) Eden -- full compliance without coercion. While all of these qualify as equilibria, all might also be unstable. In terms of the first equilibrium (E_1) it is highly unlikely that a political system would continue to tolerate noncompliance with important laws, unless the political system had been overwhelmed by noncompliance and forced to abort its enforcement effort. Indeed, if this equilibrium _is_ stable, it signifies that a political system's effort to enforce its law has failed.

Many political systems aspire toward alternatives that lie between these three degenerative equilibria, where coercion and noncompliance are balanced in a manner that may be either durable or tenuous. These alternative equilibria are located by setting Eqs. 3 and 4 to zero ($dP_t/dt = dN_t/dt = 0$), and by rearranging the result to obtain:

$$P_t = (a/b)L - (a/b)N^* \qquad (5)$$

$$P^* = S + (d/f)MN_t - (d/f)N_t^2. \qquad (6)$$

Eq. 5 defines a straight line which yields equilibrium in N_t -- symbolized as N^*. Eq. 6 defines a quadratic which yields equilibrium in P_t -- defined as P^*. Both equations are defined in the same plane, and any point of intersection satisfies the equilibrium conditions for both N^* and P^*. That is, a point of intersection yields a system equilibrium.

In order to maintain the interpretation of the model, all parameters must be positively valued and this in turn maintains system motion in the first (positively valued) quadrant of the N,P plane -- the only quadrant that provides meaningful values for the system states. This is no assurance, however, that an intersection actually occurs within the first quadrant. Indeed, the scenarios which follow show that zero, one, or two such intersections may occur, and these intersections may either produce (locally) stable equilibria, or thresholds which demarcate stability from instability.

The implication is clear: if an intersection does not occur then a nonzero equilibrium is nonexistent and either noncompliance or coercion will be driven from the system. Alternatively, if an intersection does occur, it may or may not provide a stable balance between coercion and noncompliance.

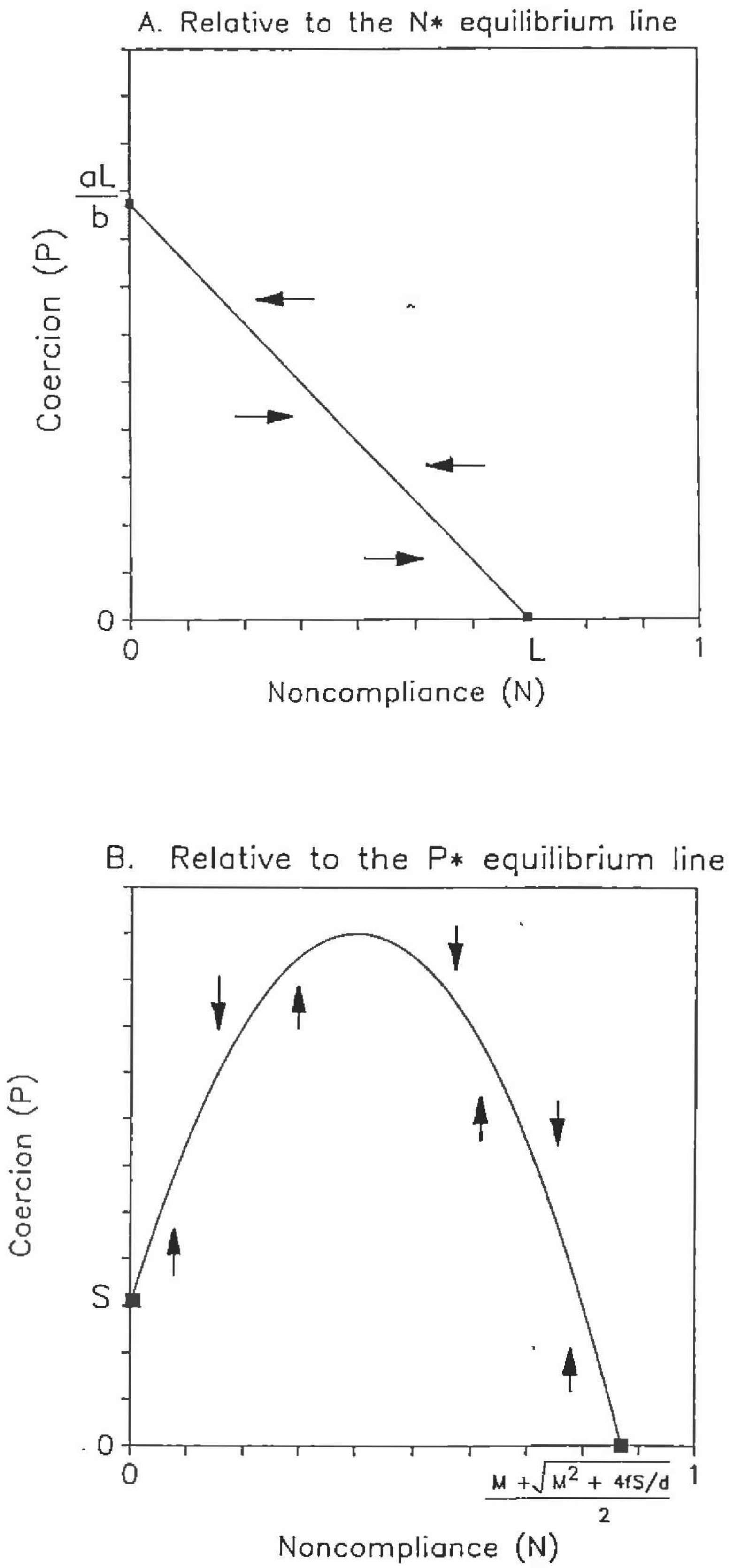
Thus it is important to develop a technology that allows a determination of system motion in the N,P plane (Maynard Smith, 1968; Huckfeldt, Kohfeld, and Likens, 1982). Figure 1a graphs the N^* equilibrium line of Eq. 5, and Figure 1b graphs the P* equilibrium line of Eq. 6. The trajectories shown in Figure 1a illustrate that first quadrant movement in N_t is always in the direction of the N* equilibrium line, and Figure 1b illustrates that first quadrant movement in P_t is always in the direction of the P* equilibrium line. Thus, simultaneous motion in P and N can be determined by taking account of both the vertical and horizontal "pulls" of each equilibrium line. In this manner the process can be allowed to unfold in time, where time is represented as a path in the N,P plane. (See the appendix for a more systematic development.)

Several other features of Figure 1 require attention because they serve as the basis for the analysis that follows. The N* equilibrium line intersects the P axis at: (a/b)L, and it intersects the N axis at: L. The P* equilibrium line intersects the P axis at: S. It intersects the N axis at two points, but only one point lies above zero: $[M + \sqrt{M^2 + 4fS/d}\,]/2$. Finally, the P* equilibrium line reaches its maximum in P when: N=M/2. These features of the equilibrium lines allow the development of several different scenarios, and the derivation of several results regarding the dynamics of coercion and noncompliance. The first scenario is concerned with the potential success of preemptive strategies in efforts to control noncompliance.

The Potential for Preemptive Success

This analysis begins by considering two scenarios, both of which might result in a successful preemptive strategy for eliminating noncompliance. In one scenario the preemptive strategy always succeeds but in the other it might fail, for reasons that are not entirely obvious. These two scenarios are shown in Figure 2. In Part A of the Figure there is no intersection between the two equilibrium lines. Thus a nonzero equilibrium is

Figure 1. Movement in the N,P plane.



not present, and motion in the system is toward: $E_1=(0,S)$. That is, noncompliance is driven out of the system, and coercion is maintained at its preemptive, fair share equilibrium. An intersection is present in Part B of the figure, but it is unstable, producing a threshold effect. Levels of noncompliance and coercion within a particular subarea of the plane produce runaway noncompliance and the demise of the coercive effort. Levels of noncompliance which lie outside this area produce movement toward E_1, and the elimination of noncompliance. The alternatives in this second scenario are either the complete elimination of noncompliance or runaway noncompliance, and only slight differences in initial conditions or minor exogenous shocks to the system might determine the result.

What do the two scenarios share in common? In both instances:

$$S > (a/b)L. \quad (7)$$

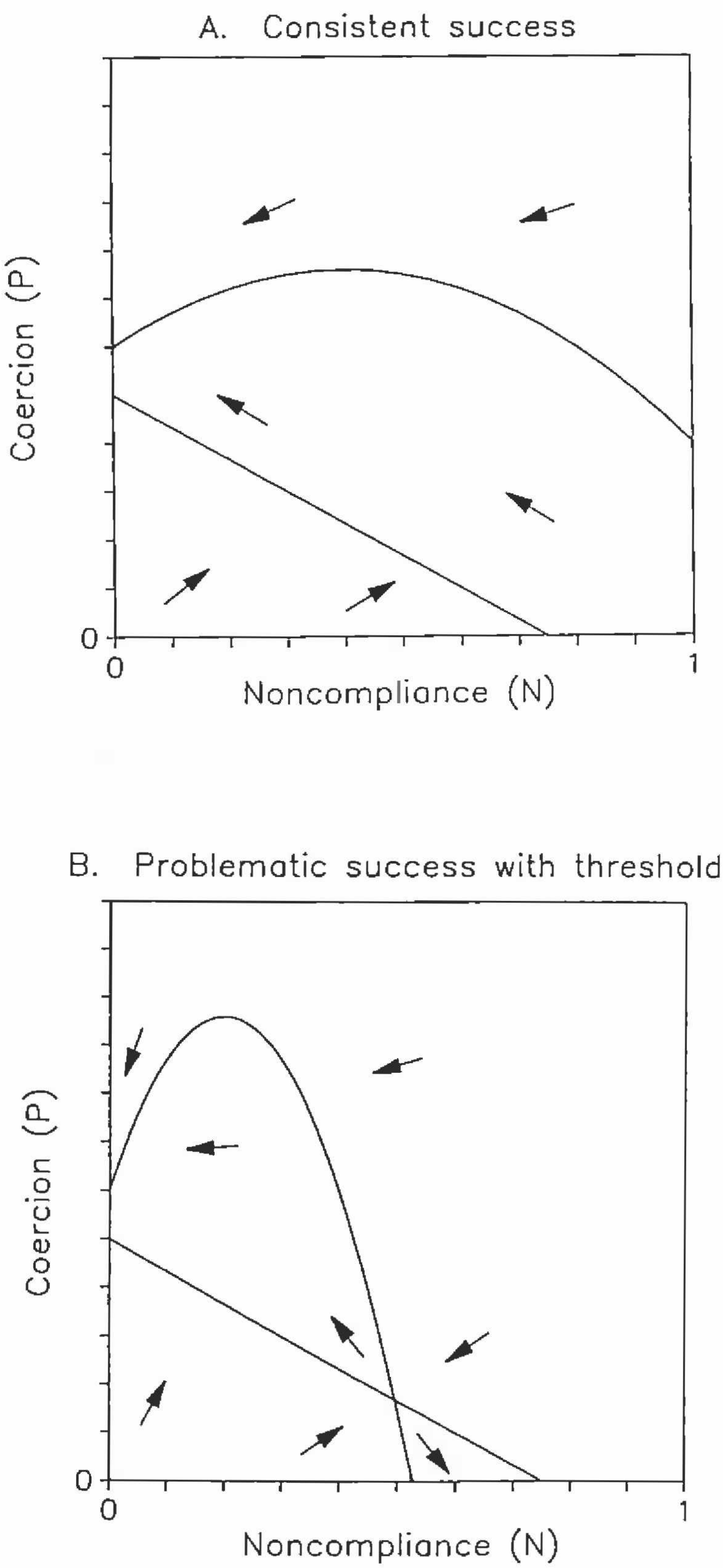
And it can be shown that this condition must hold if noncompliance is to be completely eliminated. This means that a political system is more likely to drive out noncompliance by (1) investing heavily in preemptive coercion (S), and (2) increasing the effectiveness of its enforcement activity (b). The potential for eliminating noncompliance is also furthered by: (1) a smaller proportion of citizens who are potentially willing to break the law (L), and (2) a lower level of profit (a) from noncompliance.

What is different about the two scenarios? In the first,

$$[M + \sqrt{M^2 + 4fS/d}\,]/2 > L. \quad (8)$$

In the second scenario this inequality is reversed. First notice that Ineq. 8 <u>must</u> hold if M is greater than L -- if the noncompliance saturation point lies beyond the potential level of noncompliance. Thus, as legitimacy increases (as L decreases), and as the ability to withstand noncompliance increases, so does the potential for eliminating noncompliance. Indeed, a guarantee of

Figure 2. Preemptive strategies.



total compliance can be written wholly in terms of: preemptive coercion, coercion effectiveness, noncompliance profitability, system legitimacy, and the system's noncompliance saturation point. This guarantee leaves out any consideration of the political system's short-term response to noncompliance (f and d).

Can the preemptive strategy be guaranteed to succeed (can Ineq. 8 be satisfied) if potential noncompliance (L) exceeds the noncompliance saturation point (M)? Possibly, but herein lies an irony: the potential for consistently eliminating noncompliance is furthered by exercising restraint in the short-term response to noncompliance! Notice that Ineq. 8 is more likely to be satisfied if f grows in size and d declines in size. But this means that the preemptive strategy is more likely to succeed if the political system is unwilling or unable to increase its level of coercive resources above its fair share level, and if the system is less responsive to noncompliance in the short-term. (For interesting and complementary results generated using a different model and different assumptions see Salert and Sprague, 1980.)

Why? The exercise of coercion is risky business because it has inherent built-in limitations. Political systems that are willing and able to respond aggressively to noncompliance run the inherent risk of making themselves more vulnerable to the noncompliance saturation point. In graphic terms, the P^* line declines more rapidly after reaching its maximum, and thus it is more likely to intersect with the N^* line, producing a threshold.

In summary, the preemptive strategy is more likely to succeed in eradicating noncompliance if the political system invests heavily in a sustained fair share for the coercive apparatus, but then exercises restraint in its short-term response to noncompliance. Not only can such a political system afford patience, but patience pays its own dividends.

Controlling Noncompliance without Preemption

A successful preemptive strategy imposes potentially significant costs, especially in political systems with low levels of legitimacy (high L), and where the profitability of noncompliance (a) is high. First, the burden of a large resource share for the coercive apparatus must be maintained. Second, and perhaps more important, many people do not find it congenial to live in a police state, even if the coercive apparatus does exercise restraint in its response. The question then arises: what are the prospects for controlling noncompliance, absent a preemptive strategy?

The absence of a preemptive strategy may be represented formally by reversing Ineq. 7 so that:

$$(a/b)L > S. \quad (9)$$

In words, the preemptive strategy is abortive if the fair share for coercion is less than potential noncompliance weighted by the ratio of noncompliant profitability to coercion effectiveness. Statements such as these are, of course, relative. It is possible to pursue a preemptive strategy with only a small level of S, so long as L is also very small. Conversely, it may not be possible to pursue the preemptive strategy even if S is very large, depending upon the magnitude of L.

The absence of a preemptive strategy has at least two consequences that not a few law abiding citizens might find congenial: (1) lower taxes, or more resources for other functions of the political system, and (2) the potential for a more open society. Is such a society doomed to rampant lawlessness?

Two scenarios, both of which either forsake or fail to attain the preemptive strategy, are shown in Figure 3. In Part A of the figure a stable equilibrium is attained between coercion and noncompliance. Any displacement away from this equilibrium is self-extinguishing. While noncompliance is not driven out of the system, it is contained at a level somewhere between complete compliance ($N^*=0$), and the maximum possible level of noncompliance ($N^*=L$).

In contrast, Part B of Figure 3 shows runaway noncompliance and the demise of the coercive effort. Noncompliance reaches its maximum level and coercive resources go to zero. At best (or worst) this signifies that the political system has given up enforcing an unpopular law. At worst (or best) it signifies the complete unravelling of the political system.

Ineq. 9 holds for both scenarios. How are the scenarios different? In the first instance of the stable equilibrium, where noncompliance is successfully controlled, it is the case that:

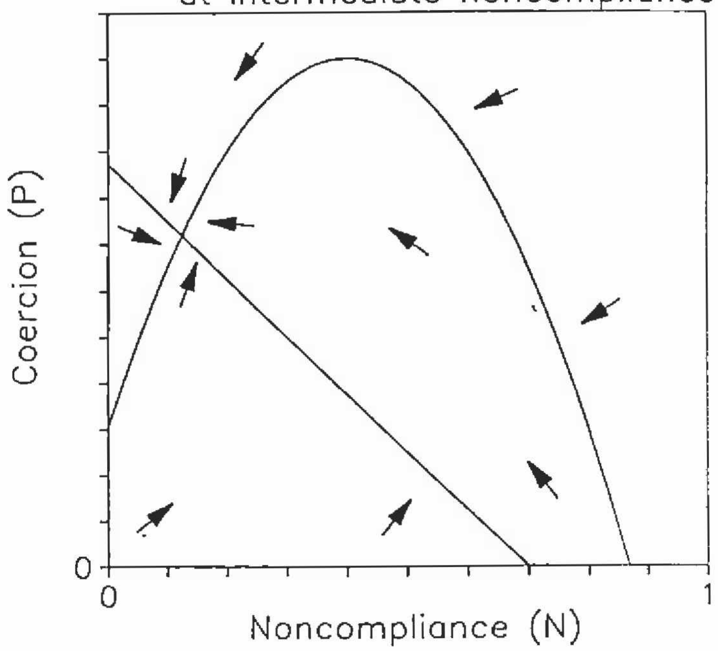
$$[M + \sqrt{M^2 + 4fS/d}\,]/2 > L. \qquad (10)$$

A reversal of Ineq. 10 will not necessarily produce the pattern of runaway compliance shown in Figure 3b; an alternative outcome, shown below in Figure 4, might also be obtained. For present purposes it is sufficient to observe that Ineq. 10 guarantees a stable equilibrium.

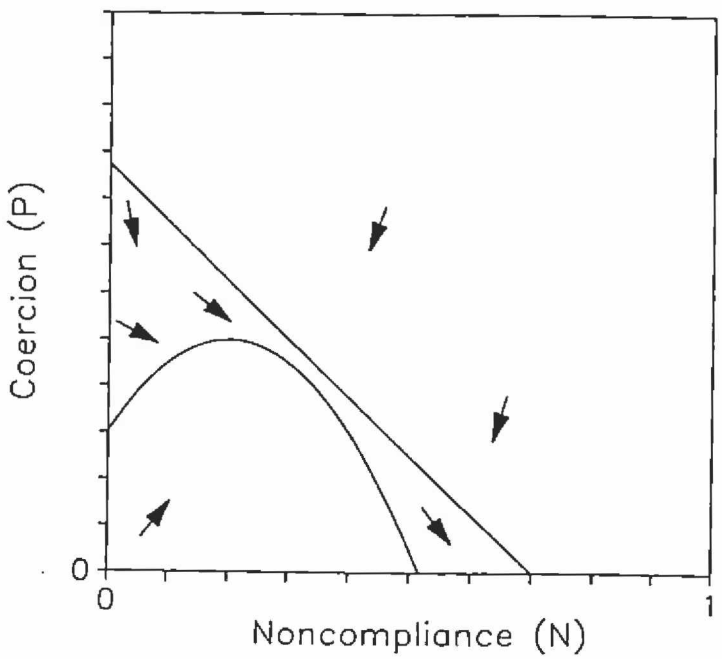
Recall that Ineq. 10 is the same condition that separated the consistently successful preemptive strategy from the preemptive strategy where success was both problematical and fragile. As before, success is guaranteed if $M > L$: if the point of noncompliance saturation is greater than the highest possible level of noncompliance, thereby making saturation impossible. And as before, restraint in the short-term response to noncompliance is the best policy. Even if $L > M$, restraint in the short run may be sufficient to generate stability. In mathematical terms, such restraint may be sufficient to satisfy Ineq. 10.

Figure 3. Non-preemptive strategies.

A. Consistent success at intermediate noncompliance



B. Consistent failure



The important point is as follows: if a system's point of noncompliance saturation (M) exceeds the highest possible level of noncompliance that might be experienced (L), then either a preemptive or a non-preemptive strategy will be successful. The preemptive strategy will drive noncompliance out of the system, and the non-preemptive strategy will control noncompliance at some intermediate level. If this inequality is reversed, if L exceeds M, then the political system is generally well advised to exercise great restraint in its response to noncompliance. The dynamic logic of both scenarios suggests that runaway noncompliance might be avoided if the short-term response of the political system is restrained. An aggressive response only serves to make the political system more vulnerable to the noncompliance saturation point.

Desperate Strategies for Desperate Circumstances

Does it ever make good sense for the political system to react aggressively to noncompliance? To continue with the scenario of Figure 3b, suppose that the potential for noncompliance (L) is so much larger than the noncompliance saturation point (M) that no amount of patience and restraint can remedy the situation and bring noncompliance under control. What is a police chief to do?

First notice that this situation (M<<L) suggests a political system with profound difficulties. It means that either (1) legitimacy is very low, or that (2) the system is very vulnerable to even minor levels of noncompliance. Fortunately or unfortunately, such situations are not rare events because vulnerability and legitimacy are likely to be inversely correlated. As citizens become increasingly willing to break the law, they

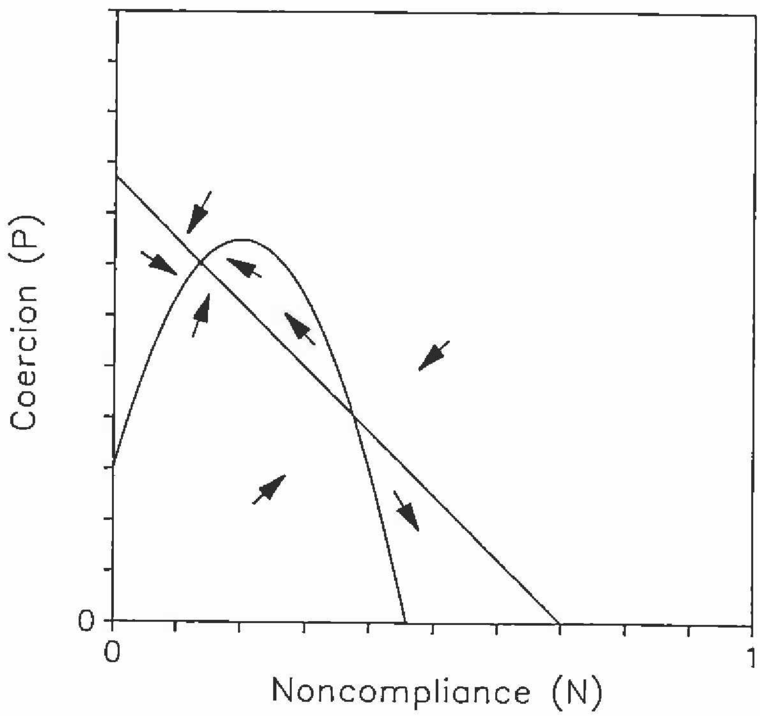
also become less willing to see the political system enforce it. Thus, as L increases, M decreases, and vice versa.

These considerations lead to a final scenario, depicted in Figure 4. The two inequalities that generated Figure 3b hold for this figure as well. In graphic terms, the difference between the two figures lies in the height of the hump for the P* equilibrium line. In Figure 3b its maximum lies below the N* equilibrium line, but in Figure 4 its maximum lies above the N* equilibrium line, and thus two equilibria are produced.

As Figure 4 shows, the left hand side equilibrium is locally stable, so that displacements in its neighborhood produce movement back toward equilibrium. The right hand side equilibrium is inherently unstable: any displacement away from this equilibrium leads either to the left hand side equilibrium or to the demise of coercion and runaway noncompliance. Thus, the second equilibrium produces a threshold that lies between successful control of noncompliance on the one hand and system degeneration on the other. In this case, as in the threshold phenomenon of Figure 2b, initial conditions become profoundly important. For example, if a new law is passed, initial levels of compliance may very well determine whether the law can be successfully enforced, or whether runaway noncompliance is the inevitable result (Likens and Kohfeld, 1983). Similarly, initial levels of enforcement activity may very well determine whether the enforcement effort succeeds or fails.

Furthermore, these models are deterministic, but reality includes a stochastic component that produces random variation in time paths. Thus, a deterministic trajectory that is headed toward a stable equilibrium may be pushed over the threshold by a random event, thereby leading to runaway noncompliance and failure on the part of the enforcement effort. In summary,

Figure 4. Non-preemptive strategy: Problematic success with threshold.



while the political system of Figure 4 is better situated than the political system of Figure 3b, the advantage may not be enormous, and the degenerative outcome may be the same. (For several interesting discussions of threshold phenomena see Schelling 1978.)

What separates Figure 4 from Figure 3b? A double equilibrium must be produced if the maximum value of P for the P* equilibrium line lies above the N* equilibrium line along a line drawn perpendicular to the N axis. This will be the case if:

$$S + dM^2/4f > a(2L-M)/2b \qquad (11)$$

The left hand side of Ineq. 11 is the maximum value of P obtained by the P* equilibrium line, where the maximum in P is obtained at: N=M/2. The right hand side of Ineq. 11 is the value of P obtained by the N* equilibrium line at the same value of N -- M/2. Ineq. 11 is <u>not</u> the general condition for the existence of intersections between the two equilibrium lines. It is possible for intersections to occur even when Ineq. 11 is not satisfied. However, as L becomes increasingly larger than M -- as circumstances become more desperate for a political system -- it becomes increasingly unlikely that intersections will occur unless Ineq. 11 is satisfied.

Several features of Ineq. 11 merit attention. First, notice that this condition reverses many of the conditions that produced the present predicament. For example, higher levels of preemptive coercion help to satisfy the inequality, but a low level of preemptive coercion relative to system legitimacy, is largely responsible for producing this desperate scenario. All of which is to indicate that, by definition, this condition is only likely to be satisfied in a political system where legitimacy is <u>very</u> low, and levels of preemptive coercion are correspondingly high -- a political system that is both desperate <u>and</u> repressive. Second, notice that Ineq. 11 is more likely to be satisfied by increasing d and reducing f: by responding aggressively to noncompliance in the short-term, and by allowing flexibility in the fair share for coercive resources.

In short, the last hope for desperate leaders in desperate circumstances is to respond in an aggressive fashion to noncompliance, both in the long-term and in the short-term. But the benefits of such a strategy must be seen within the desperate context that made it beneficial, and the tenuous situation that is produced even if the strategy is successful.

Conclusion

This effort has drawn a distinction between long-term and short-term responses to noncompliance. A political system's level of preemptive coercion -- coercive resources in the absence of noncompliance -- constitutes a long-term response, either to noncompliance, or to the potential for noncompliance. It represents the enduring commitment that a political system makes, independent of particular noncompliance levels, to coercion as a device for securing compliance. The resulting magnitude of the coercive apparatus must be acknowledged as a central feature of any society. In this sense it represents a central institutional ingredient within political systems.

This long-term response is usefully contrasted to a short-term response: the actual level of resources committed to enforcement activities in response to changing levels of noncompliance. Short-term responsiveness depends both upon the fiscal capacities of the political system, and upon system aggressiveness in responding to noncompliance. The short-term response is likely to be most pronounced for systems that are: (1) willing and able to exercise flexibility in adjusting coercive resources, and (2) aggressive in response to increased levels of noncompliance.

The analysis undertaken here produces a nonobvious result: short-term restraint coupled with long-term aggressiveness offers the best potential for eliminating noncompliance. That is, only a preemptive strategy can succeed in eradicating noncompliance, and such a strategy is furthered by exercising short-term restraint in response to noncompliance. Short-term aggressiveness has the unintended consequence of making the political system more vulnerable to its point of noncompliance saturation, and hence more likely to be overwhelmed by noncompliance.

Furthermore, even in situations where a preemptive strategy is not pursued, short-term restraint will generally afford the best strategy for controlling noncompliance at some intermediate level. Only under the most desperate circumstances, in situations where a political system is both vulnerable to noncompliance and suffering from a low level of legitimacy, do increased levels of short-term aggressiveness pay any dividends. And even in these situations, the success of the strategy is problematic.

Finally, under some circumstances short-term restraint is <u>guaranteed</u> to result in the successful control of noncompliance. In political systems distinguished by high levels of legitimacy and by an ability to withstand high levels of noncompliance, short-term restraint is certain to keep noncompliance within manageable bounds, regardless of whether a system pursues a preemptive or nonpreemptive long-term strategy.

In summary, this analysis has important consequences for democratic prospects. Not only does it suggest that short-term coercive restraint is likely to pay practical dividends, but it also suggests that noncompliance with unpopular laws might be controlled without a fundamental institutional committment to excessive levels of preemptive coercion. The success of such restraint depends, inevitably, upon the system's level of legitimacy. But increasing legitimacy is not an impossible task for democratic systems. Indeed, it should be one of the things they do best.

Appendix

This appendix gives more detailed attention to Figure 1, the graphical framework for determining motion in the N,P plane. Each equilibrium line demonstrates the set of points at which a system state, considered separately, must be at rest. Taken together, any intersection provides a point where both states must be at rest.

Define two functions R_p and R_n, and recall that coercion is at equilibrium when:

$$R_p = gS - gP + dMN - dN^2 = 0,$$

and noncompliance is at equilibrium is at equilibrium when:

$$R_n = aL - aN - bP = 0.$$

These functions do not only determine equilibria, however, they also determine the direction of change in the respective system states. When R_p or R_n is less than zero, a decrease is produced in the corresponding system state. When R_p or R_n is greater than zero an increase is produced in the corresponding system state.

In the discussion that follows, each equilibrium line is at first considered separately to determine the effect of deviations from the equilibrium line upon one dimensional movement in the phase plane. The area above and to the right of the noncompliance equilibrium line contains points that are higher on either P or N, thereby producing a negative R_n and a decrease in noncompliance. Correspondingly, the area below and to the left contains points that are lower on either P or N, thereby producing a positive R_n and an increase in noncompliance. Thus, horizontal movement in the N,P plane -- change in the level of noncompliance -- is attracted toward the noncompliance equilibrium line.

Calculating movement in the vertical direction -- change in noncompliance -- is more complex due to the nonlinearity of the coercion equilibrium line. As a first step consider change in R_p with respect to N:

$$\partial R_p/\partial N = dM - 2dN.$$

From this it follows that the effect of N is fundamentally transformed at: N=M/2. Increased levels of N below the maximum contribute toward positive change in P, and decreased levels contribute toward negative change in P. Conversely, increased levels of N above the maximum contribute toward negative change in P, and decreased levels contribute toward positive change. In contrast to the effect of N, increased values for P always contribute toward negative change in N, and decreased levels always contribute toward positive change.

As a second step it is useful to divide the P* equilibrium line vertically at its maximum: N=M/2. Now consider the area above the equilibrium line and to the left of the its maximum. This area contains points that are either higher on P or lower on N than the points on the equilibrium line, and this produces a decrease in P. The area above the equilibrium line and to the right of the maximum contains points that are either higher on P or higher on N than the points on the equilibrium line, and this also produces a decrease in P. The area below the equilibrium line and to the left of the maximum contains points that are either lower on P or higher on N, producing an increase in P. And finally the area below the equilibrium line and to the right of the maximum contains points that are either lower on P or lower on N, also producing an increase in P. Taken together, this all means that vertical motion above the coercion equilibrium line is downward, and vertical motion below the coercion equilibrium line is upward.

Any point on the plane is subject to a horizontal attraction and a vertical attraction. The strength of each attraction is a direct function of the distance from the respective equilibrium line. If a point is removed from both equilibrium lines, its instantaneous motion will generally lie in a diagonal direction. If a point lies at an intersection, it will not move at all. If a point lies on one equilibrium line but not on another, its instantaneous motion will either be at an angle of 0 or 90 degrees, depending upon which equilibrium line it is located on. This one-directional motion will only last for an instant, however, as the two-dimensional attraction is renewed as soon as the trajectory departs from the equilibrium line.

The analysis of this paper is not, in general, concerned with the nature of movement toward (or away from) equilibrium. In particular, the possibility of limit cycles is not explicitly considered, even though it would appear that the system considered here does not generate such motion. For discussions of these issues see May (1974).

References

Barry, Brian. M. 1970. Sociologists, Economists and Democracy. London: Collier-Macmillan.

Becker, Gary. S. 1968. Crime and Punishment: An Economic Approach, Journal of Political Economy, 76: 169-217.

Cortes, Fernando, Adam Przeworski and John Sprague. 1974. Systems Analysis for Social Scientists. New York: Wiley.

DeNardo, James. 1985. Power in Numbers: The Political Strategy of Protest and Rebellion. Princeton: Princeton University Press.

Gillespie, John V., Dina A. Zinnes, Philip A. Schrodt, G. S. Tahim, and R. Michael Rubison. 1977. An Optimal Control Model of Arms Races, American Political Science Review, 71: 226-244.

Gilpin, Michael E. 1975. Group Selection in Predator-Prey Communities. Princeton: Princeton University Press.

Granovetter, Mark. 1978. Threshold Models of Collective Behavior, American Journal of Sociology, 83: 1420-1443.

Gurr, Ted Robert. 1970. Why Men Rebel. Princeton: Princeton University Press.

Huckfeldt, R. Robert. 1983. The Social Context of Political Change: Durability, Volatility, and Social Influence, American Political Science Review, 77: 929-944.

Huckfeldt, R. Robert, C. W. Kohfeld, and Thomas W. Likens. 1982. Dynamic Modeling: An Introduction. Beverly Hills: Sage.

Kuklinski, James H. and John E. Stanga. 1979. Political Participation and Government Responsiveness: The Behavior of California Superior Courts, American Political Science Review, 73: 1090-1099.

Likens, Thomas W. and C. W. Kohfeld. 1983. Models of Mass Compliance: Contextual or Economic Approach? Political Methodology, 9: 377-406.

May, Robert M. 1974. Stability and Complexity in Model Ecosystems. Princeton: Princeton University Press.

Monas, Sidney. 1961. The Third Section: Police and Society in Russia under Nicholas I. Cambridge: Harvard University Press.

Olson, Mancur. 1971. The Logic of Collective Action: Public Goods and the Theory of Groups. Revised Edition. New York: Schocken Books.

Przeworski, Adam. 1974. Contextual Models of Political Behavior, Political Methodology, 1: 27-61.

Rosenzweig, Michael L. 1969. Why the Prey Curve has a Hump, American Naturalist, 103: 81-87.

Rosenzweig, M. L. and R. H. MacArthur. 1963. Graphical Representation and Stability Conditions of Predator-Prey Interaction, American Naturalist, 97: 209-223.

Richardson, Lewis F. 1960. Arms and Insecurity. Pittsburgh: Boxwood.

Salert, Barbara. 1976. Revolutions and Revolutionaries: Four Theories. New York: Elsevier.

Salert, Barbara and John Sprague. 1980. The Dynamics of Riots. Ann Arbor: Interuniversity Consortium for Political and Social Research.

Schelling, Thomas C. 1978. Micromotives and Macrobehavior. New York: Norton.

Smith, John Maynard. 1968. Mathematical Ideas in Biology. Cambridge: Cambridge University Press.

Smith, John Maynard. 1974. Models in Ecology. Cambridge: Cambridge University Press.

Tilly, Charles. 1978. From Mobilization to Revolution. Reading, Massachusetts: Addison-Wesley.

Wildavsky, Aaron. 1974. The Politics of the Budgetary Process. Second Edition. Boston: Little, Brown.