Endogenous Nonconvex State-Transition Rules and Cyclical Policies

C. Y. Cyrus Chu

SFI WORKING PAPER: 1991-08-031

SFI Working Papers contain accounts of scientific work of the author(s) and do not necessarily represent the views of the Santa Fe Institute. We accept papers intended for publication in peer-reviewed journals or proceedings volumes, but not papers that have already appeared in print. Except for papers by our external faculty, papers must be based on work done at SFI, inspired by an invited visit to or collaboration at SFI, or funded by an SFI grant.

©NOTICE: This working paper is included by permission of the contributing author(s) as a means to ensure timely distribution of the scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the author(s). It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author’s copyright. These works may be reposted only with the explicit permission of the copyright holder.

www.santafe.edu
ENDOGENOUS NONCONVEX STATE-TRANSITION RULES
AND CYCLICAL POLICIES

C. Y. Cyrus Chu

Institute of Economics
Academica Sinica
Nanking, Taipei, TAIWAN, R.O.C.

May, 1990
Abstract

We show in this paper that when one's micro decision to abide by or break the law is affected by the attitude of other potential violaters, the dynamic path of macro law-breaking rate will usually be nonconvex. We also show that, because of this nonconvexity, stationary law enforcement policies may be dominated by oscillatory ones, and this result provides a possible justification for short-lived enthusiasm on the part of the law enforcement agency. A further implication of the above result is that the traditional discussion of optimal enforcement under a time-stationary setup, which excludes cyclical policies \textit{a priori}, may be misleading.

Acknowledgments

This research is part of the Economics Research Program at the Santa Fe Institute which is funded by grants from Citicorp/Citibank and the Russell Sage Foundation and by grants to SFI from the John D. and Catherine T. MacArthur Foundation, the National Science Foundation (PHY-8714918) and the U.S. Department of Energy (ER-FG05-88ER25054). I wish to thank Professors W. Brian Arthur, Michele Boldrin, John Miller, and participants of the 1989 Santa Fe Summer program for their very helpful comments and suggestions.
Endogenous Nonconvex State-Transition Rules
and
Cyclical Policies
by
C.Y. Cyrus Chu

I. Introduction

In Figure 1, we depict the monthly data of the number of traffic violations in Taipei from 1977 to 1989. Besides the slight growth pattern overall, we observe several obvious peaks in 1977, 1980, 1982, 1984, 1987 and 1988, representing the city government’s six "sweeping traffic rectification" (STR) campaigns. In these STR periods, traffic law enforcement was tightened, and the number of traffic tickets issued often doubled or tripled. As the probability of getting a ticket increases evidently in these periods, rash drivers are deterred, and traffic usually improves significantly. However, these STRs are meant to be short-term campaigns, and after the heat cooled down, traffic gradually worsened, until the city government feels enough public pressure to start a new STR campaign.

[Insert Figure 1 About Here]

The city government is often criticized for having only short-lived enthusiasm, but the fact is with extremely high costs associated with the tough enforcement in STR periods, the city government simply can not afford them as long-term programs. Indeed, when enforcement is tough, although the social cost of traffic order is low, the enforcement cost is high; whereas when enforcement is lax, although the cost of social order is high, the enforcement cost is low. This situation is not dissimilar to the negative relationship between enforcement cost and social order cost associated with most illegal or externality-creating activities, and such negative relationship has prompted some law economists suggest that there may be an "optimal" enforcement level somewhere in between the two extremes that can minimize total social costs. In the case of traffic law enforcement; what we are interested to know is if there is also an optimal enforcement level somewhere between the easy going and the stringent?

Intuitively, oscillatory policies such as the above-mentioned transitory STRs can never be optimal if the decision maker is typically assumed to minimize a convex social cost function over a convex feasible set. Indeed, as the convexity in question refers to
variables indexed by time, cycles could not be optimal, because averaging over these cycles will exploit the convexity of the function and therefore reduce the cost value achieved. The purpose of this paper is to show that, when the macro traffic order (the state variable) is formed by interactions among micro individual drivers, it is not unusual to have a nonconvex dynamic transition rule for the state variable, and hence short-lived enthusiasm (which represents a kind of cyclical enforcement policies) may actually be better than stationary policies. A more general implication is that when one's decision to abide by or break the law is affected by the attitude of other potential violaters, the discussion of optimal enforcement under a time-stationary setup, which excludes cyclical policies a priori, may be misleading.

The structure of this paper is arranged as follows. We shall present in section II a model which characterizes how individual drivers affect, and at the same time are affected by, the attitude of other drivers. It will be shown that the dynamic traffic order transition rule would naturally have a nonconvex shape. In section III we shall discuss the impact of various traffic law enforcement policies on traffic order equilibria. It will be shown that under some parametric specifications, every stationary policy can be dominated by some oscillatory ones, and this result provides a possible justification for short-lived enthusiasm on the part of the government. The fourth section elaborates the discussion in previous sections, and presents some possible future research directions.

II. A Model of Traffic Order Evolution

It has been argued in various economic literature that many types of macrobehavior can be explained as outcomes of interactions among individuals. In what follows we shall interpret the macrobehavior as the societal traffic order, and propose a model to characterize the mutual influences between individuals' micro decisions and the macro traffic order. The analysis to be presented bears some resemblance to the models in Arthur (1988) and Schotter (1978), and can be treated as a formalization of the story given in Schelling (1978).

We shall consider a population with very large size $N$. Each member of the population randomly meets (one by one, say, at the crossroads) $n$ others in each period. Because no one is in a position to know in advance whom he is going to come across, any ex-ante cooperative negotiation is out of the question. The interactions
between any two individuals will be characterized by the following symmetric 2 \times 2 game. There are two choices in everyone’s strategy space: 1 and 2. Let us interpret strategy 1 as driving rudely and strategy 2 as driving courteously. The ideal situation is for both sides to drive courteously \((\max\{a, b, c, d\} = d)\). If one expects the other driver to drive rudely, he would incline to requite like for like \((a > c)\). Conversely, if one drives rudely but the driver he comes across happens to be courteous, he would be mortified for what he has done \((d > b)\). 4

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(a, a)</td>
<td>(b, c)</td>
</tr>
<tr>
<td>2</td>
<td>(c, b)</td>
<td>(d, d)</td>
</tr>
</tbody>
</table>

Row chooser’s payoff listed first.

There are three obvious Nash equilibria in this 2 \times 2 game: \((1, 1)\), \((0, 0)\), and \((x, x)\), where the first (second) argument in the parenthesis represents the probability that the first (second) player takes strategy 1, and

\[
0 < x \equiv \frac{d-b}{a-c+d-b} < 1
\]

Since one does not know in advance what strategies the others are going to adopt, he can only draw on his experience and make an educated guess. Suppose there are \(p_t\) proportion of people taking strategy 1 (and hence \(1 - p_t\) percent taking strategy 2) at period \(t\), and in each period one comes across \(n \ll N\) people on the street randomly. The number of people taking strategy 1 (denoted \(k\)) in this \(n\)-person sample will follow a binomial distribution:

\[
k \sim \binom{n}{k} p_t^k (1 - p_t)^{n-k}
\]

Let \(\bar{Z}_n\) be the sample mean of this \(n\)-sample, then we have

\[
p_r(\bar{Z}_n > x \mid p_t) = p_r(n\bar{Z}_n > nx \mid p_t)
\]

\[
= 1 - \sum_{y=0}^{x^*} \binom{n}{y} p_t^y (1 - p_t)^{n-y} \equiv f(p_t; x)
\]

where \(x^*\) is the largest integer no greater than \(nx\). Clearly, when \(p_t = 0\), the event \(\bar{Z}_n > x\) never happens, and hence \(f(0; x) = 0\). Similarly, \(f(1; x) = 1\).

For the time being let us suppose people are myopic in the sense that they use the sample mean they observe at period \(t\) to predict the proportion of strategy-1 choosers.
in period $t + 1$. This is consistent with the setup in Schelling (1978). In section IV we shall discuss the more elaborated case where people use Bayes' formula to adjust their prior belief. Given $p_t$, it is shown in (1) that a particular individual $i$ will have probability $f(p_t; x)$ to observe a sample mean (denoted $\bar{Z}_n^i$) larger than $x$. Let

$$\omega^i = \begin{cases} 1, & \text{if } \bar{Z}_n^i > x; \\ 0, & \text{if } \bar{Z}_n^i \leq x. \end{cases}$$

Suppose we pick a sample of size $M$ and calculate the corresponding sample mean $\bar{\omega}_M = \sum_{i=1}^{M} \omega^i / M$, then clearly $\bar{\omega}_M$ will have a sampling distribution with mean $f(p_t; x)$. By Khintchine's theorem, as $M \to \infty$, $\bar{\omega}_M$ converges to $f(p_t; x)$ in probability. Thus, in a city with very large number of drivers, there will be $f(p_t; x)$ proportion of them whose $\omega^i$ equals 1 (or who observe $\bar{Z}_n^i > x$). Under our myopic prediction assumption, person $i$ with $\omega^i = 1$ would expect to meet more than $x$ proportion of strategy-1 takers in period $t + 1$, and therefore, according to the specification of the $2 \times 2$ traffic interaction game, $i$'s best response is to adopt strategy 1 in period $t + 1$. As such, in a large city, the dynamic evolution of $p_t$ will be characterized by the following equation:

$$p_{t+1} = f(p_t; x)$$

There are several possibilities for the shape of $f(p_t; x)$. First, if $0 < nx < 1$, then $x^* = 0$, and $f(p_t; x) = 1 - (1 - p_t)^n$. This implies that $df/dp_t = n(1 - p_t)^{n-1} > 0$, and $d^2f/dp_t^2 = -n(n - 1)(1 - p_t)^{n-2} < 0$. The second possibility is when $n - 1 < nx < n$, and the $f$ function becomes $f(p_t; x) = p_t^n$. Both these cases are depicted in Figure 2. As we shall see in the next section, the value of $x$ will be determined by the level of government's traffic law enforcement, and the above-mentioned two cases will stand only if the enforcement is extremely lax or tough. In what follows, we shall set aside these two extremes and concentrate on the third possibility: $1 \leq nx \leq n - 1$, where we can rewrite $f(p_t; x)$ as

$$f(p_t; x) = 1 - (1 - p_t)^n - \sum_{y=1}^{x^*} \binom{n}{y} p_t^y (1 - p_t)^{n-y}$$

It is shown in the Appendix that the $f$ curve is uniformly increasing in $p_t$, and has a unique inflexion point at $x^*/(n - 1)$. It is also clear that the $f$ curve is neither
concave nor convex. Given that $f(0;.) = 0$ and $f(1;.) = 1$, there are clearly three steady states for $p_t$ (see Figure 2).

Although there are several models explaining the interdependency of individuals' choices, our analysis is nevertheless unique. By assuming that people make their decisions on the basis of the binomial sample mean they previously observed, we were able to derive the exact formula, and the exact inflection point of the state transition rule. Thus, instead of proposing vaguely that there may be multiple equilibria as in Kuran (1987) or Arthur (1988), we can conclude that as long as $1 \leq nx \leq n-1$ holds, there are exactly three equilibria, as shown in Figure 2. This property is very helpful in developing our later discussion of government policies.

Figure 2 shows that the dynamic evolution of $p_t$ satisfies the four properties of a complex system characterized in Arthur (1988) and David (1988), namely, possible multiple equilibria, possible inefficiency, lock-in, and path dependence. The lock-in property of the dynamic evolution in Arthur and David limits the scope of policy analysis because there is hardly anything any government can do to extricate itself from a stable steady state.7

III. Welfare Impact of Different Enforcement Policies

To make the model more realistic for policy analysis, let us assume that there are $q_1$ proportion of (well-mannered) people who never drive rudely, and $q_2$ proportion of (rash) people who always drive rudely. With these modifications, the dynamic transition rule of the state variable $p_t$ becomes:

$$p_{t+1} = q_2 + (q_1 - q_2)f(p_t; x) \equiv g(p_t; x)$$

and is shown in Figure 3, with the number of steady states being possibly one, two or three.

Suppose a penalty of $\pi$ dollars is mandatory on all detected rude drivers. Suppose the probability of detection is $\tau$, then the expected penalty would be $p \equiv \tau \cdot \pi$. Taking
into account the possible penalty associated with adopting strategy 1, the payoff matrix of the 2 × 2 game should be modified as:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$a - \rho, a - \rho$</td>
<td>$b - \rho, c$</td>
</tr>
<tr>
<td>2</td>
<td>$c, b - \rho$</td>
<td>$d, d$</td>
</tr>
</tbody>
</table>

The corresponding critical value would now become

$$
\bar{x} \equiv \frac{d - (b - \rho)}{a - \rho - c + d - (b - \rho)} = \frac{d - b + \rho}{a - c + d - b} = \frac{d - b + r\pi}{a - c + d - b} \quad (5)
$$

and people would take strategy 1 if they expect more than $\bar{x}$ proportion of others to take it too. Let $\bar{x}^*$ be the largest integer no greater than $n\bar{x}$, then the dynamic transition rule of $p_t$ should now become

$$
p_{t+1} = g(p_t; \bar{x}) = q_2 + (1 - q_1 - q_2)[1 - \sum_{y=0}^{\bar{x}^*} \binom{n}{y} p_t^y(1 - p_t)^{n-y}] \quad (6)
$$

Given fixed fine, (5) tells us that tightened enforcement (increasing $r$) would increase $\bar{x}$. If this increase is significant enough to cause the threshold integer $\bar{x}^*$ to go upward, then individuals' behavior would change, the curve $g(p_t; \bar{x})$ would shift down, and the corresponding steady state would also change. Let $c(r, p^*(r))$ denote the total (social order and enforcement) costs associated with $r$ and its corresponding steady-state traffic order $p^*(r)$. The conventional argument is: whether a change in $r$ is worth doing depends upon the sign of $\Delta c/\Delta r$.

Suppose we are now at point A of Figure 3 which is sustained by very strict traffic law enforcement with cost too high to afford persistently. Suppose the government is trying to relax the enforcement in order to cut cost. As $r$ reduces, $\bar{x}$ reduces, the curve $g(p_t; \bar{x})$ shifts up, and the steady state value $p^*$ increase to $B$. However, the nonconvexity of $g$ implies that as $r$ continues to reduce there may be a jump in $p^*(r)$, and hence a jump in social costs. This could be seen in Figure 3 where a slight reduction of $r$ at the point $C$ would make the steady state $p$ value change from $p_c$ to $p_D$. Similarly, there would also be a jump if one tries to increase $r$ at the point $E$. In fact, no $p$ value between $p_E^*$ and $p_c^*$ could ever be sustained as a steady state. Thus, for stationary policies, the government is faced with only 2 choices: very good traffic
order sustained by strict enforcement (AC area), or very bad traffic order as a result of lax enforcement (EF area).

Instead of focusing on stationary policies, we may take a look at oscillatory policies and find out if they fare any better. This has been made possible by the fact that with the intrinsic nonconvexity of \( g(\cdot, \cdot) \), averaging over two points of a cycle may end up at a point which can not be sustained by any stationary policies. Let us assume that violation fines collected by the government are paid back to the public in lump-sum so that they are private but not social cost. Thus, if we normalize the population size to be one, the expected social return of traffic interactions in period \( t \) will be \( a p_t + (b + c)p_t(1 - p_t) + d(1 - p_t)^2 \equiv B_t \). Enforcement cost \( C_t \) is assumed to be a linear function of detection probability: \( C_t = \alpha \cdot r_t \), and hence the net social benefit (or negative social cost) would be

\[
W_t \equiv B_t - C_t = ap_t^2 + (b + c)p_t(1 - p_t) + d(1 - p_t)^2 - \alpha r_t \equiv W(p_t, r_t)
\]

where the relationship between \( r \) and \( p_{t+1} \) is governed by (5) and (6). With \( \delta \) the social discount rate, the present value of total social benefit is \( \sum_{t=0}^{\infty} \delta^t W_t \).

In the following numerical calculation, we consider six stationary enforcement policies (detection probabilities): \( r = (r^0, ..., r^5) = \{0, 0.1, 0.2, 0.3, 0.4, 0.5\} \), and the corresponding \( x \)'s will be \( x^i = (d - b + r^i \pi)/(a - c + d - b) \). Let \( p^*(i) \) be the steady state(s) achieved when \( r = r^i \), it can be shown that when \( i = 0, 3, 4, 5 \), \( p^*(i) \) has only one value, whereas when \( i = 1, 2 \), \( p^*(i) \) has two possible values. In the latter case we shall list both of them in the Tables. Total discounted social benefit under stationary enforcement policy \( r^i \) will be

\[
\sum_{t=0}^{\infty} \delta^t W_t = \sum_{i=0}^{\infty} \delta^i W(p^*(i), r^i) \equiv W(p^*(i), r^i) = \frac{W(p^*(i), r^i)}{1 - \delta} \equiv TW_{sp}^i
\]

where "sp" stands for stationary policy.

Now consider the following cyclical policy which shuttles between \( r^0 \) and \( r^5 \). Suppose at period zero \( p_0 = p^*(0) \) which may be very large because there is essentially no enforcement. Suppose the government decides that starting from period one, \( r \) will be increased to \( r^5 \), a very strict enforcement. Then \( p_t \) will follow the path \( p_{t+1} = g(p_t, x^5) \), and the social benefit in each period will be \( W_t = W(p_t, r^5) \). The state variable \( p_t \) eventually converges to \( p^*(5) \), (see Figure 4). After staying at \( p^*(5) \)
for a while, suppose at period $T_j$, the government decides to relax enforcement to $r_0$ from period $T_j + 1$ onward. The $p_t$ will then follow the path $g(p_t; \tilde{x}^0)$, and the social benefit in each period is $W_t = W(p_t, r^0)$. The government can repeat the above-mentioned cycles once $p_t$ converges to $p^*(0)$. This cyclical enforcement policy would generate a sequence of $W_t$'s, and the discounted total social benefit will be denoted $TW_{cycle}^0$, where the superscript "0" specifies the cycle's starting point $p^*(0)$. Similarly, if at period zero $p_0 = p^*(5)$ which is very small, we assume that the government start the cyclical policy by first reducing $r$ to $r^0$, then as $p_t$ converges to $p^*(0)$, increasing $r$ to $r^5$, and so on. Discounted total social benefit so obtained will be denoted as $TW_{cycle}^5$. For starting points $p_0 = p^*(i), i = 1, 2, 3, 4$, the government will first adopt $r^0(r^5)$ if $p^*(i)$ is small (large), and will begin to switch policies (between $r^0$ and $r^5$) afterwards. We shall demonstrate below that under some parametric specifications, $TW_{sp}^i < TW_{cycle}^i$ for all $i$, i.e., the government will have no incentive to stay at any stationary state.

We assume $a = 4, b = 8, c = 0, d = 10, \pi = 5, \alpha \in \{0, .5, 1.0\}$, and $\delta \in \{.95, .90, .85, .80\}$. For each of the 12 cases, we calculate $p^*(i)$, $TW_{sp}^i$ and $TW_{cycle}^i$ respectively, and the results are summarized in Tables 1 - 3. When $\alpha = 0$, traffic law enforcement is costless, and the government will have no incentive to leave a steady state with good traffic order (e.g., $p^*(5)$). Under these circumstance, short-lived enthusiasm cannot be justified. But when $\alpha = 1$ (Table 3), $p^*(5)$ starts to lose its attractiveness because the enforcement cost to sustain $p^*(5)$ is too high. At any time $t$ with $p_t = p^*(5)$, the government can always realize higher total social benefit by cyclical enforcement swings between $r^5$ and $r^0$. The same also holds true for $p_t = p^*(i), i = 0, 1, 2, 3, 4$. This implies that all these stationary policies can be dominated by the $r^0 - r^5$ cyclical policy, and thus short-lived enthusiasm is justified.

In Table 2 where $\alpha = .5$, the situation is more interesting. It can be seen that when $\delta = .95$, staying at the second and the fourth stationary state will result in higher social benefit, and hence the government will have no reason to make any policy change. But when $\delta$ reduces to .80, every stationary policy will be dominated by the corresponding oscillatory one, and cyclical policies become better. It is interesting to note that the relationship between the desirability of cycles and the size of the
discount rate is consistent with the findings in the literature of chaotic dynamics of Boldrin (1988).

IV. Discussion and Extensions

1. Bayesian Decision Makers

In the previous two sections we assume that people are myopic in the sense that they use the sample mean they observe at period $t$ to predict the proportion of strategy-1 takers in period $t + 1$. It would be interesting to see whether our derived result will still hold if individual drivers use a more complex Bayesian formula to make their predictions.\(^\text{10}\)

Suppose individual $i$ at period $t$ has a prior belief $f_i^t(\cdot)$ about the distribution of the true proportion of strategy-1 drivers in the society. The sampling distribution of observing $k$ strategy-1 drivers out of the $n$-sample, given that $\pi$ is the societal mean, is

\[
p(k \mid \pi) = \binom{n}{k} \pi^k (1 - \pi)^{n-k},
\]

Thus, for person $i$ who observes $k$ out of the $n$-sample at period $t$, his posterior should be adjusted as

\[
f_{t+1}^i(\pi \mid k) = \frac{\binom{n}{k} \pi^k (1 - \pi)^{n-k} \cdot f_i^t(\pi)}{\int \binom{n}{k} \pi^k (1 - \pi)^{n-k} \cdot f_i^t(\pi) d\pi}
\]

Suppose person $i$ decides to take strategy 1 but the true $\pi$ turns out to be $\pi < p^*$, then his loss will be $[a\pi + b(1 - \pi)] - [c\pi + d(1 - \pi)] \equiv L_1(\pi)$. On the other hand if person $i$ chooses strategy 2 but $\pi > p^*$, then his loss will be $[c\pi + d(1 - \pi)] - [a\pi + b(1 - \pi)] \equiv L_2(\pi)$. The difference between the posterior expected losses of adopting strategy 1 and strategy 2 will be

\[
D \equiv \int_0^{p^*} f_{t+1}^i(\pi \mid k) L_1(\pi) d\pi - \int_{p^*}^1 f_{t+1}^i(\pi \mid k) L_2(\pi) d\pi,
\]

which can be simplified as

\[
D = (d - b) - (a - c + d - b) \int_0^1 f_{t+1}^i(\pi \mid k) \pi d\pi \quad (8)
\]
Let $\int_0^1 f_{i+1}(\pi | k) \pi d\pi$ be denoted $\hat{\pi}_{i+1}$, then it is clear from (8) that person $i$ would adopt strategy 1 if and only if

$$\hat{\mu}_{i+1} > \frac{d - b}{a - c + d - b} = p^*$$

The total number of strategy-1 takers at period $t + 1$ (denoted $N_{i+1}^t$) can then be counted as $i$ goes through 1, 2, ..., $N$, and the true value of $\pi$ at period $t$ will be $N_{i+1}^t/N$. By taking the Dirichlet distribution (which is the conjugate family of Bayesian binomial sampling) to run a simulation, we have found that both the nonconvexity of $g(\cdot, \cdot)$ and the possible dominance of cyclical policies still hold under this Bayesian setup.\footnote{\textsuperscript{11}}

2. Cycles in Other Illegal Activities

It is often argued that one's decision to abide by or break the law depends on the attitude of other potential violaters. Indeed, as Lui (1986) pointed out, if most of one's neighbours evade taxes, or most of one's colleagues are corrupt, his decision to do the same is always made after taking into consideration the cost of peer group condemnation, and the net of mutual coverage. This kind of interaction among people also tends to make the dynamic path of macro crime index nonconvex. Furthermore, such nonconvexity may be helpful in explaining the persistence of crime waves, which will be explained below.

As Cooter and Ulen (1988) and Wilson (1983) pointed out, U.S. crime statistics over the last decades showed that the amount of a wide variety of crimes declined from a peak in the mid-1930s to a low point in the early 1960; then began a rapid increase from the early 1960s till the mid or late 1970s; and has begun to decline slowly in the 1980s. There are two often-mentioned explanations for these crime cycles: the first relates the peak of the crime rate to the possible unfair distribution of income in the period of rapid economic growth; the second hypothesis suggests that crime cycles may be related to the cyclical age structure brought by birth waves.\footnote{\textsuperscript{12}} What was not stressed in previous discussion of this topic was the interaction between crime waves and enforcement policies. When the crime rate is high, there usually exists public pressure asking for tightening the law enforcement in order to improve social order, as described in Cooter and Ulen (1988, p. 534). Thus, as the crime rate reach its peak, there is a natural tendency for it to go down. Furthermore, since the dynamic time path of crime rates is nonconvex, a tightened law enforcement may cause a persistent
fall of crime rates, as shown by the zigzag curve from F to A in Figure 4. Similarly, when the crime rate is low, people may also propose to relax the law enforcement in order to reduce the seemingly unnecessary enforcement costs, and the proposed lax enforcement could also cause a persistent increase in crime rates, as shown from A to F. As such, aside from the exogenous demographic or economic shocks, the above-mentioned public pressure together with the nonconvexity of dynamic path form a endogenous force of prolonging crime cycles. This mutual interaction between crime rates and enforcement policy seems to be an important factor in interpreting empirical data, and perhaps one can set up a Granger (1969) causality model to refine the existing empirical literature in the future.

3. Optimal Non-stationary Policies

We have demonstrated in section III that under some parametric specifications any stationary policy can be dominated by a cyclical one. One implication of this result is that if the government intends to search for an "optimal" enforcement policy, the traditional time-invariant setup is intrinsically misleading. With a nonconvex dynamic state transition rule (6), the government’s problem would become

$$\max_{r_t} \sum_{t=0}^{\infty} \delta^t W(p_t, r_t)$$

s.t.  $$\ddot{x}_t = \frac{d - b + r_t \pi}{a - c + d - b}$$

$$p_{t+1} = g(p_t; \ddot{x}_t)$$

(9)

The problem in (9) will generate an optimal sequence $r_0^*, r_1^*, ..., r_t^*, ...$, which may or may not converge to a stationary value. It would be worth studying what exactly the necessary or sufficient conditions for (12) to have a "complex" enforcement dynamics are, and the paper by Brock (1988) may be a good starting point for future research along this line.
Appendix

Differentiating the R.H.S. of (3) with respect to \( p_t \) yields

\[
\frac{df}{dp_t} = n(1 - p_t)^{n-1} - \sum_{y=1}^{x^*} \binom{n}{y} [y p_t^{y-1}(1 - p_t)^{n-y} - (n - y)p_t^n(1 - p_t)^{n-y-1}] \quad (A1)
\]

\[
= n(1 - p_t)^{n-1} - n \sum_{y=1}^{x^*} \binom{n-1}{y-1} p_t^{y-1}(1 - p_t)^{n-y} - \sum_{y=1}^{x^*} \binom{n-1}{y} p_t^n(1 - p_t)^{n-y-1}
\]

Since

\[
\binom{n-1}{y-1} p_t^{y-1}(1 - p_t)^{n-y} \bigg|_{y=k} = \binom{n-1}{y} p_t^n(1 - p_t)^{n-y-1} \bigg|_{y=k-1},
\]

by expanding the terms in the square bracket of (A1), one finds that most of them cancel with each other, and \( df/dp_t \) can be further simplified as

\[
\frac{df}{dp_t} = n(1 - p_t)^{n-1} - n[(1 - p_t)^{n-1} - \left( \frac{n-1}{x^*} \right) p_t^{x^*}(1 - p_t)^{n-x^*-1}]
\]

\[
= n \left( \frac{n-1}{x^*} \right) p_t^{x^*}(1 - p_t)^{n-x^*-1} > 0 \quad (A2)
\]

From (A2), one can easily derive

\[
\frac{d^2 f}{dp_t^2} = n \left( \frac{n-1}{x^*} \right) p_t^{x^*-1}(1 - p_t)^{n-x^*-2}[x^*(1 - p_t) - (n - x^* - 1)p_t]
\]

\[
= n \left( \frac{n-1}{x^*} \right) p_t^{x^*-1}(1 - p_t)^{n-x^*-1}[x^* - p_t(n - 1)], \quad (A3)
\]

which will be positive (negative) if \( p_t \) is less (larger) than \( x^*/(n - 1) \).
References


Footnotes

1. The literature was initiated by the seminal paper of Becker (1968). Later extension in various directions can be found in, e.g., Stern (1978), Polinsky and Shavell (1979), Posner (1985), and Pyle (1983).

2. See Boldrin (1988) for detailed explanation.

3. See Jones (1976), Schelling (1978), Lui (1986), and Arthur (1988) for discussion of such interaction in various areas of economics.

4. The setup here is in accordance with the "critical mass" model in Chapter 3 of Schelling (1978).

5. In his discussion about the number of seminar participants, Schelling (1978, p. 105) assumes that people use the number who attended last week as an expectation of attendance this week.


7. The only exception is in Arthur (1988) where he made a short discussion about the possibility of "shaking" the system into new configurations so as to leave an inferior local steady state.

8. It is noticed in equations (5) and (6) that a change in r could effectively shift the curve \( g(p_i; \bar{x}) \) only when this change is significant enough to alter the threshold integer \( \bar{x}^* \). Thus, although \( r \) may have infinitely many values, under a given \( \pi \) only a finite number of them could make \( \bar{x}(r) \equiv d - b + r\pi/(a - c + d - b) \) integers. As such, for welfare analysis we only have to consider finite \( r \)'s with their corresponding \( \bar{x} \)'s integers, because any \( r \) with \( \bar{x}(r) \) between two neighbouring integers \( (m < \bar{x}(r) < n) \) will have the same enforcement impact as \( r_m \), but have higher enforcement cost than \( r_m \), where \( r_m \equiv \bar{x}^{-1}(m) \). This is why we can, without loss of generality, consider only finite cases.

9. Policy options should also include cases where \( r > 5 \). But readers should be able to verify that our conclusions still hold in these cases.

10. I am indebted to Luca Anderlini for his comments and suggestion on this issue.

11. Detailed discussion about the Bayesian approach is available from the author. Properties of the conjugate Dirichlet family can be found in DeGroot (1970).

12. However, a closer examination shows that these two explanations seem insufficient to account for the crime waves we observe, see Cooter and Ulen (1988) Chapter 12 for explanations.
Table 1: Net social benefit under stationary and cyclical enforcement policies ($\alpha = 0$)

<table>
<thead>
<tr>
<th>$i$</th>
<th>$p^*(i)$</th>
<th>$TW^i_{sp}$</th>
<th>$TW^i_{cycle}$</th>
<th>$TW^i_{sp}$</th>
<th>$TW^i_{cycle}$</th>
<th>$TW^i_{sp}$</th>
<th>$TW^i_{cycle}$</th>
<th>$TW^i_{cycle}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.800</td>
<td>84.72</td>
<td>125.47</td>
<td>42.41</td>
<td>63.10</td>
<td>28.27</td>
<td>41.84</td>
<td>21.20</td>
</tr>
<tr>
<td>1</td>
<td>.217</td>
<td>153.51</td>
<td>124.93</td>
<td>76.76</td>
<td>63.26</td>
<td>51.17</td>
<td>42.72</td>
<td>38.38</td>
</tr>
<tr>
<td>2</td>
<td>.797</td>
<td>84.92</td>
<td>125.51</td>
<td>42.45</td>
<td>63.14</td>
<td>28.31</td>
<td>41.88</td>
<td>21.23</td>
</tr>
<tr>
<td>3</td>
<td>.203</td>
<td>153.32</td>
<td>125.96</td>
<td>78.16</td>
<td>64.14</td>
<td>52.11</td>
<td>43.47</td>
<td>39.08</td>
</tr>
<tr>
<td>4</td>
<td>.200</td>
<td>156.79</td>
<td>126.11</td>
<td>78.40</td>
<td>64.28</td>
<td>52.26</td>
<td>43.59</td>
<td>39.20</td>
</tr>
<tr>
<td>5</td>
<td>.200</td>
<td>156.80</td>
<td>126.11</td>
<td>78.40</td>
<td>64.28</td>
<td>52.27</td>
<td>43.59</td>
<td>39.20</td>
</tr>
</tbody>
</table>

Table 2: Net social benefit under stationary and cyclical enforcement policies ($\alpha = .5$)

<table>
<thead>
<tr>
<th>$i$</th>
<th>$p^*(i)$</th>
<th>$TW^i_{sp}$</th>
<th>$TW^i_{cycle}$</th>
<th>$TW^i_{sp}$</th>
<th>$TW^i_{cycle}$</th>
<th>$TW^i_{cycle}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.799</td>
<td>84.81</td>
<td>145.35</td>
<td>42.40</td>
<td>73.09</td>
<td>28.27</td>
</tr>
<tr>
<td>1</td>
<td>.217</td>
<td>143.51</td>
<td>144.81</td>
<td>74.25</td>
<td>73.26</td>
<td>49.50</td>
</tr>
<tr>
<td>2</td>
<td>.797</td>
<td>79.92</td>
<td>145.39</td>
<td>39.96</td>
<td>73.13</td>
<td>26.64</td>
</tr>
<tr>
<td>3</td>
<td>.202</td>
<td>146.31</td>
<td>145.83</td>
<td>73.15</td>
<td>74.13</td>
<td>48.77</td>
</tr>
<tr>
<td>4</td>
<td>.200</td>
<td>136.79</td>
<td>146.00</td>
<td>63.39</td>
<td>74.27</td>
<td>45.59</td>
</tr>
<tr>
<td>5</td>
<td>.200</td>
<td>131.79</td>
<td>146.00</td>
<td>63.39</td>
<td>74.27</td>
<td>43.93</td>
</tr>
</tbody>
</table>

Table 3: Net social benefit under stationary and cyclical enforcement policies ($\alpha = 1.0$)

<table>
<thead>
<tr>
<th>$i$</th>
<th>$p^*(i)$</th>
<th>$TW^i_{sp}$</th>
<th>$TW^i_{cycle}$</th>
<th>$TW^i_{sp}$</th>
<th>$TW^i_{cycle}$</th>
<th>$TW^i_{cycle}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.799</td>
<td>84.81</td>
<td>165.24</td>
<td>42.40</td>
<td>83.09</td>
<td>28.27</td>
</tr>
<tr>
<td>1</td>
<td>.217</td>
<td>143.51</td>
<td>164.70</td>
<td>71.75</td>
<td>83.26</td>
<td>47.83</td>
</tr>
<tr>
<td>2</td>
<td>.797</td>
<td>74.92</td>
<td>165.28</td>
<td>37.46</td>
<td>83.13</td>
<td>24.97</td>
</tr>
<tr>
<td>3</td>
<td>.202</td>
<td>136.31</td>
<td>165.72</td>
<td>68.15</td>
<td>84.13</td>
<td>45.43</td>
</tr>
<tr>
<td>4</td>
<td>.200</td>
<td>126.73</td>
<td>165.86</td>
<td>63.36</td>
<td>84.26</td>
<td>42.24</td>
</tr>
<tr>
<td>5</td>
<td>.200</td>
<td>106.79</td>
<td>165.89</td>
<td>53.39</td>
<td>84.28</td>
<td>35.59</td>
</tr>
</tbody>
</table>
Figure 1: Number of traffic violations in Taiwan, 1977-1989.

Source: Major Statistics, City Government of Taipei

Figure 2: The Shape of $f(p_t; x)$
Figure 3: Points between $P_E^*$ and $P_C^*$ can not be sustained by stationary policies.

Figure 4: Cyclical enforcement – shuttling between $r^0$ and $r^5$. 