Continuous punishment and the potential of gentle rule enforcement

Ido Erev
Faculty of Industrial Engineering and Management
Technion
Haifa, Israel 32000

Paul Ingram
712 Uris Hall
Columbia Business School
Columbia University
New York, NY
10027

Ornit Raz and Dror Shany
Faculty of Industrial Engineering and Management
Technion
Haifa, Israel 32000

December 15, 2009

This research was supported by the Davidson Research Fund, and by a grant from Israel-USA Binational Science Foundation. The order of the authors is alphabetic, all contributed equally. We are very grateful for helpful conversations with and comments from Greg Barron, Amar Bhide Al Goldberg, Daniel Gopher, Ray Horton, James Kitts, Tamar Landesberg, Dov Monderer, Dov Zohar, and participants in the Workshop on Network Dynamics at the Russell Sage Foundation.

Continuous punishment and the potential of gentle rule enforcement

The paper explores the conditions that determine the effect of rule enforcement policies that imply an attempt to punish all the visible violations of the rule. We start with a simple game theoretic analysis that highlights the value of gentle COntinuous Punishment (gentle COP) policies. If the subjects of the rule are rational, gentle COP can eliminate violations even when the rule enforcer has limited resources. The second part of the paper uses simulations to examine the robustness of gentle COP policies to likely deviations from rationality. The results suggest that when the probability of detecting violations is sufficiently high, gentle COP policies can be effective even when the subjects of the rule are boundedly rational adaptive learners. The paper concludes with experimental studies that clarify the value of gentle COP policies in the lab, and in attempt to eliminate cheating in exams.

Keywords: The economics of small decisions, decisions from experience, underweighting of rare events, social enforcement.

1. Introduction

The “quality of life” policing strategy, implemented in New York City in the mid1990’s (see Kelling and Sousa, 2001), implies a “continuous punishment” reinforcement schedule. It advocates increased number of policemen on the streets and an attempt to punish all visible offenses. Many believe that the implementation of this strategy is responsible for the decline in crime and disorder documented in New York during the 1990’s (see Golub et al., 2002; Kelling and Sousa, 2001; Silverman, 1999). However, other researchers are skeptical. One source of skepticism is the observation that there are other explanations for the decline in crime rate in the 1990’s (see Eck and Maguire, 2000). In addition, field studies (see Harcourt 2001) question the robustness of the predicted effects. The most important debate involves the relationship between disorder and crime. The quality of life strategy rests on the assumption that disorder increases crime. This assumption was supported by Skogan (1990) who found that robbery victimization was higher in neighborhoods characterized by disorder, but Harcourt (2001) found that the crime-disorder relationship did not hold for other crime types including burglary, assault, rape and pick-pocketing.

The current paper uses basic learning research in an attempt to clarify the boundary conditions for the success of “continuous punishment” rule enforcement policies. The reliance on learning research reveals that there are many settings in which continuous punishment policies are likely to be counterproductive. Basic learning research shows that punishments tend to be less effective than positive reinforcements (see Azrin and Holz, 1966), and that continuous reinforcements tend to be less robust than partial reinforcements (see Humphreys, 1939; Ferster and Skinner, 1957).[1] Thus, understanding the exact boundary conditions for the effective implementation of continuous punishment policies is important.

The present analysis rests on the assumption that the net effect of learning is a result of an interaction between the incentive structure and the basic properties of the

[1] However, the generalization from this "partial reinforcements extinction effect" (PREE) is not immediate. Most demonstrations of the PREE focus on the effect of positive reinforcements on behavior during extinction. The current analysis focuses on the effect of punishments. In addition, the PREE is interesting, but unique: most studies reveal behavioral momentum: An increase in behavioral tendencies with the likelihood of reinforcements (see Nevin & Grace, 2001).

learning processes. Thus, we model learning within a game theoretic model of the environment. Our analysis focuses on an environment in which a rule enforcer with limited resources tries to eliminate violations of a specific rule. One example (and the focus of the empirical study described below) is a proctor that tries to eliminate cheating in an exam. We start with an easy example: A simplified setting in which a set of continuous punishment policies, referred to as "gentle COntinuous Punishment" (gentle COP), is extremely effective. Under these policies the enforcer is committed to punishing the initial violations of the rule. As a result, rational individuals prefer not to be the first violators, and the risk of extinction is eliminated.

The second stage of our analysis uses a simulation to explore the robustness of gentle COP policies to limitations of the rational game-theoretic approach. Specifically, we show that gentle COP policies can be effective even if the agents are adaptive learners. To simulate the behavior of adaptive learners we used the sampling and weighting model that was found to provide useful predictions of behavior in experimental study of decisions from experience (see Erev and Haruvy, 2009). In addition, the results reveal clear boundary conditions. The effectiveness of gentle COP policies is highly sensitive to the probability of detection of each violation.

The third and final step of our theoretical analysis examines the robustness of gentle COP policies to the limitations of the game-theoretic approach identified in sociological research of norm enforcement by social groups. Gentle COP policies designed to benefit the group can be enhanced by social sanctions. Indeed, gentle COP policies can be effectively implemented by a third party through formal sanctions, and/or by members of a group through social sanctions.

We then move on to two empirical examinations of the gentle COP idea. The first is a laboratory study, which exhibits the characteristics of limited information, and third-party sanctioning. The second is a field-study of cheating in exams, where enforcement comes both from a proctor, and from the students themselves.

1.1 A simple game-theoretic model and two classes of rule enforcement policies

The current analysis focuses on rule enforcement problems that satisfy two conditions. First, the rule enforcement unit (e.g., the proctor during an exam) has limited capacity: The potential number of violators of the rule (N) is larger than the maximal number of individuals that can be punished (M). Second, obeying the rule is costly (the cost is $C > 0$), but the fine for punished violators is larger (the fine is $F > C$).

It is convenient to distinguish between two classes of policies that can be adopted by the rule enforcement unit to eliminate the motivation to violate the rule. The first implies a uniform enforcement. Under these policies the enforcement unit punishes each of the violators with equal probability. Since the number of potential violators is N, the probability of punishment will be between M/N (when all N the potential violators violate the rule) to 1 (when the number of violators is not larger than M). It is easy to see that the effect of the uniform policies depends on the magnitude of the fine. When the fine is relatively large (when $F \geq C(N/M)$), rational agents (assuming that the payoffs represent utility) are expected to obey the rule. With lower fines (when $C < F < C(N/M)$), the uniform policies create a game with two Nash equilibrium points in pure strategy.[2]

A second class includes gentle COntinuous Punisment (gentle COP) policies. Under these policies the unit declares that the *first* M violators will be punished (and ensures that decisions are made sequentially with complete information). Given this declaration no one will be motivated to be the first violator even if the fine is relatively low (only slightly larger than C).[3]

1.2 Uniform and gentle COP policies under bounded rationality

The game-theoretic analysis presented above assumes that the subjects of the rule: know the incentive structure, play sequentially after observing the prior choices, make only one choice, and behave rationally. In addition, this simplified analysis assumes that

---

[2] In the first, no subject violates the rule. This is a Nash equilibrium because no subject will be motivated to deviate from this state unilaterally and violate the rule, because unitary violators are certain to be punished. In the second equilibrium all the subjects violate the rule. This state is a Nash equilibrium because the probability of punishment is low, and for that reason violation is the best strategy.

[3] When decisions are made sequentially under full information and the game is played once, the no violation equilibrium can be derived based on the assumption that players will select their dominant strategy. Finding the equilibrium in repeated play with partial information is beyond the scope of the current paper, although we examine that context in the simulations and experiment. Mainstream economic analyses of rule enforcements (e.g., Becker, 1968) tend to ignore this simple policy because they focus on situations (like tax problems) in which the monitoring cost is very high, and it is therefore difficult or impossible to commit to punishing all initial violations of a rule.

the enforcement unit can punish the first violation with certainty. In most rule enforcement problems these assumptions are unlikely to hold. In this section we explore the robustness of gentle COP policies to violations of these assumptions. We start with an analysis of an environment in which the only information available to the subjects is the outcome of their previous decisions. The rationality assumption is replaced here with the assumption that the subjects adapt to the obtained reinforcements.

The combination of low rationality and limited information leads to a pessimistic prediction concerning the effectiveness of many common patterns of rule enforcement. Two robust properties of decisions from experience imply that subjects are likely to violate rules even when observing the rules is the rational choice. One property involves underweighting of delayed outcomes (Hernstein, Lowenstein, Prelec and Vaughan, 1993). This property explains, for example, why workers fail to use ear protection even when the outcome (implicit punishment) is costly hearing loss in the long term (Zohar, 1980).

A second important property involves underweighting of punishments that are delivered with low probabilities. For example, Barron and Erev (2003) found that in a repeated choice task most people prefer a gamble that leads to a loss of 32 units with probability 0.1 (and no loss otherwise) over the economically more attractive option of a sure loss of 3 units. Moreover, the tendency to underweight low probability punishments increases when the probability of punishment decreases (Perry, Erev and Haruvy, 2002).

The attractiveness of gentle COP is enhanced with the recognition of these behavioral regularities. First, the moderate size of the punishments used under gentle COP policies facilitates immediacy. For example, if the punishment for cheating on an exam is heavy, it cannot be easily applied without a careful, costly and long process of review and appeal. However, if the punishment is a gentle warning, it can be immediate. Second, under gentle COP policies a large proportion of the initial violations are punished. Thus, subjects are less likely to experience low punishment probabilities.

To highlight the robustness of gentle COP to these properties of human decisions from experience, we start with a simulation analysis that compares the effect of uniform and gentle COP policies on virtual subjects that behave according to the sampling and weighting model (SAW, Erev and Haruvy, 2009). This model was found to capture the main properties on decisions from experience in binary tasks. It is a simplified variant of the model that won the first Technion choice prediction competition (see Erev et al., 2009). The basic assumptions of this model for the current setting are summarized below.

A1: Exploration and exploitation. The agents are assumed to consider two cognitive strategies: exploration and exploitation. Exploration implies a random choice. Exploitation implies a choice of the alternative with the highest "subjective value."

A2: The probability of exploration decreases with experience and the expected length of the experiment (T).[4] The exact probability of exploration in trial t is

$$P(Explore_t) = \varepsilon^{\frac{t-1}{t+T^{\delta}}}$$

where ε is a free parameter that captures exploration after long experience, and δ the sensitivity to the length of the experiment.

A3: Exploitation trials start with a draw of a sample of m_t past trials. The value of m_t at trial t is assumed to be randomly selected from the set {1, 2,…… κ} where κ is a free parameter. The first draw is the most recent trial. All previous trials are equally likely to be sampled (with replacement) in each of the other m_t -1 draws.

A4: Linear subjective value. The subjective value of alternative j at trial t is a weighted average of the "grand mean" ($GM_{j,t}$ – the mean payoff from j in the first t-1 trials) and the "sample mean" ($SM_{j,t}$ -- the mean payoff from k in the sample taken at trial t):

$$SV(k,t) = w(GM_{j,t}) + (1-w)(SM_{j,t})$$

[4]Implicit in A2 is the simplification assumption that the agents know the value of T. This assumption is not exactly correct, but it seems reasonable as the agents can estimate T from the know length of the experiment.

where w is a weighting parameter.

<Insert Figure 1>

Predictions. The current analysis focuses on a simulation of 200-trial repeated play of the enforcement game described in Section 1.1. At each trail of this simulation, four virtual agents (N=4) chose between obeying or violating "the rule." Violation saved two points (C= 2), but could lead to a fine of 3 points (F=3). The average number of punishments per trial could not exceed 1 (M=1). Thus, the maximal number of punishments over the 200 trials was 200. Two enforcement procedures were compared. In Condition Uniform the enforcement resources were distributed uniformly: one violator could be punished in each trial. In Condition gentle COP, all the resources were available early: That is, the first 200 violations were punished.

The subjects of the rule were virtual agents that behave in accordance with the SAW model with the parameters estimated by Erev and Haruvy to summarize the results they reviewed. The exact values are: $\varepsilon = 0.1$, $\delta = .5$, $\kappa = 9$, and $w = 0.25$. Two curves in the left panel of Figure 1 present proportion of rule violations in 10 blocks of 20 trials. The upper curve shows the proportion of violation in Condition Uniform. The probability of punishment of each violation at trial t is 1/(Number of violators at trial t). The lower curve shows a condition in which the rule-enforcement unit uses a gentle COP policy and punishes all the first 200 violations. The results show that this gentle COP policy reduces the total number of violations from more than 200 to less than 50. The observation that the number of violations is lower than 200 implies that the gentle COP policy avoids the extinction stage. As a result, the advantage of this policy increases with time.

Robustness analyses and boundary conditions. Five sets of analyses were conducted in order to evaluate the robustness and the boundary conditions of the predictions summarized above. The basis of all five analyses is the simulation with parameters described above, with variations as noted. In the first set, summarized in the second panel (labeled Robustness 1) of Table 1, we focused on the effect of the four parameters of the learning model. The results show that the value of the gentle COP policy is relatively robust to the learning parameters.[5]

<Insert Table 1>

In Robustness 2 we varied the number of punishments available per trial (the parameter M). This analysis shows that a decrease in the number of punishments can impair the value of the gentle COP policy. When the M is too low (below 0.3 in the current setting), gentle COP is not much more effective than the Uniform policy. Again, the logic here is that under gentle COP enforcements are applied upfront but limited in number. The fewer punishments that are available, the more likely that they will be exhausted before subjects conclude that violation doesn't pay.

Robustness 3 examined the role of the tradeoff between the probability of detection of each violation (P_{detect}) and the magnitude of the fine. The expected value of the fine given violation was kept at 3 in all analyses. Thus, the actual fine (if violation was detected) was 3/ P_{detect}. For example, when P_{detect}=0.1, the fine was 30. The results show that with the current parameters the effectiveness of the gentle COP policy is highly sensitive to the probability of detection. The gentle COP policy is ineffective when P_{detect} is lower than 0.30. That is, the probability of detection is much more important than the magnitude of the fine. When detection is low-probability, subjects' learning that violations lead to punishments is greatly inhibited. Thus, probability of detection is one of the most important contingencies for gentle COP effectiveness, and a key consideration when choosing between gentle COP and alternative enforcement regimes.

In two additional robustness tests we examine the effect of observational learning (learning from the outcomes obtained by others) and higher level of rationality (the possibility that some agents select the optimal best reply to the behavior. The results reveal that these factors intensify the pattern described above. Specifically, when the number of punishments is sufficient (M above 0.5), and P_{detect} is large (above 0.3) then observational learning and higher level of rationality facilitates the effect of the gentle COP policy.

---

[5] Additional analyses show that the advantage of gentle COP is also expected under learning models that assumes sequential adjustment of propensities. We have examined two examples of this class of models: The reinforcement learning model proposed by Erev & Roth (1998), and the value adjustment model proposed by Barron and Erev (2003).

1.3 The role of social norms.

Another important limitation of the game-theoretic model involves the unrealistic assumption that the subjects of the rule cannot affect each other directly. Many important rule-systems rely on endogenous enforcement by the subjects of the rules. Social enforcement questions the validity of the current game-theoretic model because it occurs even when it requires "altruistic punishments" (Yamagishi, 1986; Fehr and Gachter, 2002), and can counteract formal enforcement (as when restriction-of-output norms on the shop floor contradict management efforts to maximize individual output, [Homans, 1950]). However, when the formal enforcement is gentle, fair, and designed to help the group, social enforcement is expected to support the formal enforcement (Nee, 1998). In these cases social enforcement is naturally modeled with the assumptions that "the group" is part of the enforcement unit, and part of the punishment F is a social sanction (*e.g.*, Coleman, 1990; Heckathorn, 1990). As for receiving punishment, the key is that the normative sanction hurts the utility of those receiving it. The conditions under which this assumption is satisfied, while not trivial, are well known in the literature on norms. Essentially, normative sanctions are more effective the more utility the target derives from membership in the group (Homans, 1950).

Interestingly, gentle COP is for three reasons even more appealing in the social enforcement context than in the third-party enforcement context. Firstly, social groups are particularly able to provide the comprehensive monitoring and timely sanctions that lead more quickly to rule compliance under gentle COP for subjects who choose responses based on reinforcement learning (Robustness 5). Secondly, norms can act as a mechanism for social learning (Argote and Ingram, 2000), a means of communicating to group members the experience of others and thus promoting quicker learning under gentle COP (Robustness 3). Thirdly, the costs of severe punishments, and therefore the advantage of the modest punishments that gentle COP enables, are greater when the sanctions are social. The punishment for the violation of a norm is social, in the extreme, ostracism from the group. Because social punishments hurt the punished in much the same way as the punisher (they both suffer a reduction in the social relationship), there is a built-in constraint on the number of group members that can be punished—as the proportion of group members suffering punishment increases, so does the cost to other group members.

2. A laboratory experiment

Our first experiment was designed to evaluate a gentle COP policy in a context where subjects had a limited initial knowledge of payoffs and probabilities of punishment. It examines human behavior in an abstract replication of the computer simulation summarized in the left-hand side of Figure 1.

Eighty undergraduate students (average age 24, 50% female) served as paid participants in the experiment. They were randomly assigned to one of the two experimental conditions. Participants received a guaranteed $1.5 for participating in the experiment, and a bonus that was contingent upon performance and ranged between $2 and $7.

Participants were run in groups of four for 200 trials. In each trial, each participant was asked to select one of two unmarked buttons. A selection of the left button was treated as a decision to obey the rule. A section of the left button always paid 1 point. A selection of the right button was treated as a decision to violate the rule. It increased the gain (relative to the Left choice) by 2 points, but could lead to a punishment of 3 points.[6] Thus, the final payoff was 3 or 0 points.

Two conditions, Uniform and gentle COP, were compared. In the first condition, punishments were uniformly distributed, one per period. Thus, the probability of punishment was 1 divided by the number of violators. For example, if two participants violated the rule in a certain trial, the probability of punishment was 1/2. In the second, gentle COP condition, punishments were focused: The first 200 violations were punished. Thus, the probability of punishment of each violation under gentle COP was 1 during the initial "learning" trials, and 0 after the first 200 violations (but we predict that the number of violations will not exceeds 200).

[6] This incentive structure implies that the experiment explores the enforcement of the rule "Select Left." As explained below, we did not inform the participants that there was a rule, and/or that the experiment examines rule enforcement. Rather, our goal was to examine the effect of two rule enforcement procedures in a minimalistic environment in which the subjects of the rule have to base their decisions of the obtained reinforcements.

Participants received no prior information as to the game's payoff structure. They received individual feedback in every trial after all four players selected one of the two buttons. It included their payoff from the last choice and an update of their cumulative payoff counter, which was constantly displayed.

The results summarized in the right-hand side of Figure 1 show a large effect of the gentle COP manipulation to reduce rule violation. Over the 200 trails, the proportion of "violation" choices in Condition Uniform was 0.67 (STD= 0.24) and 0.08 (STD = 0.15) in Condition gentle COP. The difference was significant ($t[18] = 7.24$, $p< 0.01$). The average numbers of violations over the 200 trials was 536 in the uniform condition, and only 56 in the gentle COP condition. As a result, far more punishments were administrated in the uniform condition (more than 190 in all cohorts) than in the gentle COP condition (56). As predicted, the gentle COP groups did not reach the extinction stage. It is also obvious from the figure that there is a good correspondence between the results of the simulation and those of the experiment. This complements previous evidence regarding the empirical validity of the learning model upon which we built the simulation.

3. A field experiment

The sensitivity analysis reported above highlights two important environmental factors that can impair the effectiveness of gentle COP policies in natural settings. Gentle COP policies are not likely to be effective when the enforcement resources are too limited, and/or when the probability of detection is too low. The main goal of the current section is to demonstrate that there are important natural situations in which gentle COP policies can, nevertheless, be successful. We chose to focus on cheating in exams. Exam cheating is an appealing context partly because enforcement involves both a formal third party (the proctor), and social norms (students who may sanction cheaters).[7] It also demonstrates some amount of social learning, as students may observe the punishments obtained by others. Some indications for the potential value of gentle COP policies in this context come from the observation that students cheat when there are no apparent penalties (Bowers, 1964; McCabe and Trevino, 1997), and that non-gentle rules (severe

[7] Even the request from a fellow student to be quiet, during the exam, can serve as a social sanction.

punishments such as expulsion) are not likely to work because they are too costly to the enforcers (Schneider, 1999).

The experiment was conducted during final-semester exams of undergraduate courses. Traditionally, instructions for exam proctors at the university that hosted the study included the following points:

(1) The student's ID should be collected at the beginning of the exam,

(2) A map of students' seating should be prepared.

Since the collection of the ID is the first step in the construction of the map, the common interpretation of these instructions was that the map should be prepared at the beginning of the exam. Early preparation of the map distracts the proctors, and reduces the probability of gentle punishments (e.g., warning and/or relocation) at the beginning of the exam[8] (see Zohar, 1980 for elegant support of a similar argument).

To facilitate the implementation of gentle COP in experimental conditions we simply changed the second instruction to proctors:

(2e) "A map of the students seating should be prepared 50 minutes after the beginning of the exam."

Seven undergraduate courses were selected to participate in the study. In all courses the final exam was conducted in two rooms. One room was randomly assigned to the experimental (gentle COP) condition, and the second was assigned to the control condition. The only difference between the two conditions involved the timing of the preparation of the map in the instructions to the proctors. In the control group the instruction was:

(2c) "A map of the students seating should be prepared immediately after the beginning of the exam."

Notice that implicit in the current "manipulation" of the gentle COP strategy is the assumption that proctors naturally punish gently (e.g., by direct glances, verbal warning, and/or written comments) initial attempts to cheat (e.g. attempts to observe other students' exams). Thus, we predict that gentle COP will be implemented in the

[8] This problem could be solved by asking the proctors to prepare the map before the beginning of the exam. The current research does not evaluate this solution as it implies several changes (e.g., the proctors should work more time). Yet, the results presented below suggest that this solution is likely to be effective.

experimental condition, but not in the control condition (in which the initial resources will be invested in the map/large punishments).

After finishing the exam, the students were asked to complete a brief questionnaire in which they were ask to “rate the extent to which students cheated in this exam relative to other exams.”

The results presented in Table 2 show that in all seven cases the experimental manipulation decreased the perception of cheating in the exam. This trend is statistically significant ($p < .01$, Sign Test). So, cheating as perceived by exam-takers was reduced with only a small change in the practice of proctoring, and without increasing the resources for proctoring, or the punishment for cheating.

<Insert Table 2>

4. General Discussion

The current analysis sheds light on the potential value and on the limitations of enforcement policies that imply continuous punishment. The potential value is highlighted by the observation that policies that use continuous punishment can be effective even when the punishments are gentle (moderate), and the enforcer has limited resources. The effectiveness of gentle continuous punishment (gentle COP) policies was found to be robust to the sophistication of the agents. It is expected under the assumption of rational agents, and under the assumption that the agents are adaptive learners with limited understanding of the incentive structure.

Additional attractive features of gentle punishments are suggested by an attempt to relate the current results to the classical study of reinforcement schedules. It seems that the usage of gentle punishments reduces the risk associated with continuous punishment. The most important risk involves the emergence of counter-productive classical conditioning and avoidance learning (see Azrin and Holz, 1966). For example, when a child is heavily punished for a spelling error, he might associate writing with the aversive outcome, and learn to avoid writing. We believe that the use of gentle punishments reduces this risk. That is, it seems natural to assume that a gentle request to move to the first row during an exam, will not have a large effect on the probability to come to the next exam.

Another implication of the classical study of reinforcement schedule involves the partial reinforcement extinction effect (see Ferster and Skinner, 1957). This effect implies that behavior learnt under continuous schedules, is less robust to extinction. The usage of gentle punishments can address this problem as it can reduce the risk of extinction periods. When the punishments are gentle and can be based on social sanctions, extinction does not have to be an important problem.

The experimental studies, presented here, were designed to evaluate the practical implications of the theoretical analysis. They highlight examples of situations in which gentle COP policies are effective. Two interesting properties of the gentle COP solution are clarified in these studies. First, the implementation can be surprisingly easy: a slight modification of the instructions to proctors that facilitates mild initial punishments significantly reduces the perception of cheating during exams. Second, this simple solution in not always used where it would be effective. In the past, the university that hosted our study tried to eliminate cheating with the threat of (delayed and rare) severe punishments. To support this threat proctors were asked to prepare the map that hindered initial mild punishments.

The important limitation of gentle COP policies, demonstrated here, involves the effect of the probability of detection. When the probability of detection is low (below 0.3 in the current analysis), the advantage of gentle COP policies is eliminated. This observation (and see similar finding in Luria, Erev and Zohar, 2007) implies a natural explanation to the apparent inconsistency of previous studies of the effect of quality of life and similar order facilitating policies. As noted above, order was found to be associated with lower robbery victimization rate (Skogan, 1990), but not with other crime types including burglary, assault, rape and pick-pocketing (Harcourt, 2001). Under the current explanation, order facilitating policies are examples of gentle COP policies: There effectiveness depends on the probability of detecting violations. Since they imply more policemen in the streets, they are more effective in reducing street robbery (higher detection probability) than in reducing crimes like burglary, assault, rape and pick-pocketing that cannot be easily observed by the added policemen.

Table 1
Sensitivity Analyses of the gentle COP and Uniform policies
under different learning and environmental parameters

Analysis	Learning parameters							proportion of violations	
	κ	w	ε	δ	Implied model	P_{detect}	Number of punishmens per trial	Uniform	Gentle COP
Baseline from Figure 1	8	0.25	0.1	0.55		1	0	0.79	0.10
Robustness 1: The effect of learning parameters	**1**	0.25	0.1	0.55	Best reply	1		0.12	0.10
	20	0.25	0.1	0.55				0.82	0.10
	8	**0**	0.1	0.55	Sampling			0.76	0.10
	8	**1**	0.1	0.55	Fictitious play			0.79	0.10
	8	0.25	**0.01**	0.55				0.71	0.03
	8	0.25	**0.2**	0.55				0.76	0.16
	8	0.25	**0.3**	0.55				0.72	0.21
	8	0.25	0.1	**0.1**				0.64	0.06
	8	0.25	0.1	**0.9**				0.72	0.21
Robustness 2: The effect of the number of punishments per trial	8	0.25	0.1	.55		1	0.1	0.89	0.83
							0.3	0.89	0.32
							0.7	0.86	0.10
							0.9	0.84	0.10
Robustness 3: The effect of the probability of detection	20	0.3	0.01	.5	1	0.1	0	0.60	0.54
						0.3		0.62	0.29
						0.7		0.77	0.16
						0.9		0.80	0.11

Table 2

Levels of Cheating in Seven Exams

Course	Control Condition	Experimental Condition (gentle COP)
	Mean (N)	Mean (N)
1*	1.83 (12)	1.28 (14)
2	2.0 (13)	1.55 (18)
3	2.25 (24)	1.5 (27)
4	2.2 (24)	1.57 (21)
5**	1.81 (17)	1.63 (19)
6	2.58 (31)	1.84 (19)
7***	2.36 (19)	2.07 (13)

* Multiple choice

** Open book

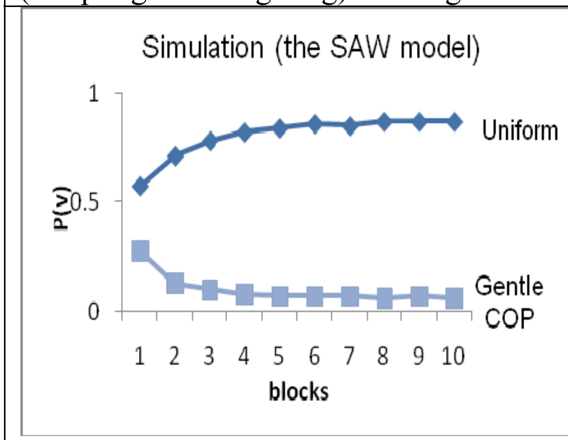
*** Students from two faculties

Figure 1: Results of the Simulation and the Laboratory Experiment

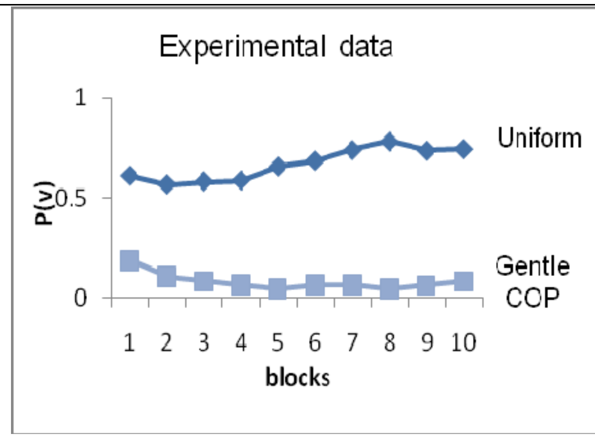
<table>
<tr>
<td>Simulation Results:

Proportion of violations in 10 blocks of 20 trials, subjects assumed to follow the SAW (sampling and weighting) learning model.</td>
<td>Laboratory Experiment Results:

Proportion of violations in 10 blocks of 20 trials</td>
</tr>
<tr>
<td>



</td>
<td>



</td>
</tr>
</table>

References

Argote, L. and Ingram, P., 2000. Knowledge transfer: the basis of competitive advantage in firms. Organizational Behavior and Human Decision Processes, 82: 150-169.

Azrin, N. H., and Holz, W. C. 1966. Punishment. In W. K. Honig (Editor), Operant behavior: Areas of research and application. New York: Appleton-Century-Crofts, pp 380-447.

Barron, G., and Erev, I., 2003. Small feedback based decisions and their limited correspondence to description based decisions. Journal of Behavioral Decision Making, 16: 215-233.

Becker, G., 1968. Crime and Punishment: An Economic Analysis. Journal of Political Economy, 76: 169-217.

Bowers, W.J., 1964. Student Dishonesty and its Control in College. New York: Columbia University (Bureau of Applied Social Research).

Coleman, J., 1990. Foundations of Social Theory. Cambridge, MA: Harvard University Press

Eck, J. and Maguire, E., 2000. Have changes in policing reduced violent crime? An assessment of the evidence. In: Alfred Blumstein and Joel Wallman (Editors.), The Crime Drop in America. New York: Cambridge University Press.

Erev, I. and Haruvy, E. in press. Learning and the economics of small decisions. In Kagel, J.H. and Roth, A.E. (Eds.), *The Handbook of Experimental Economics*. Princeton University Press. http://www.utdallas.edu/~eeh017200/papers/LearningChapter.pdf

Erev, I. and Roth, A. E., 1998. Prediction how people play games: Reinforcement learning in games with unique strategy equilibrium. American Economic Review, 88: 848-881.

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S., Hau, R. Hertwig, R. Stewart, T., West, R. and Lebiere, C., 2010. A choice prediction competition, for choices from experience and from description. Journal of Behavioral Decision Making.

Fehr, E., Fischbacher, E. and Gächter, S., 2002. Strong Reciprocity, Human Cooperation and the Enforcement of Social Norms. Human Nature, 13: 1-25.

Ferster, C. B. and Skinner, B. F., 1957. Schedules of reinforcement. New York: Appleton-Century-Crofts.

Golub, A., Johnson, B., Taylor, A. and Liberty, H., 2002. The validity of arrestee's self reports. Justice Quarterly, 19: 477-502.

Harcourt, B. E., 2001. Illusion of Order: The False Promise of Broken Windows Policing. Cambridge, Mass.: Harvard University Press.

Heckathorn, D. D., 1990. Collective sanctions and compliance norms: a formal theory of group-mediated social control. American Sociological Review, 55: 366-84

Herrnstein, R.J., Lowenstein, G. F., Prelec, D., Vaughan, W., Jr., 1993. Utility maximization and meliorization: Internalities in individual choice. Journal of Behavioral Decision Making, 6: 149-183.

Homans, G. C., 1950. The Human Group. New York: Harcourt Brace Jovanovich

Humphreys, L. G., 1939. Acquisition and extinction of verbal expectations in a situation analogous to conditioning. Journal of Experimental Psychology, 25: 294-301.

Jacob, B. and Levitt. S., 2003. Rotten apples: An investigation of the prevalence and predictors of teacher cheating. The Quarterly Journal of Economics, 118: 843-877.

Kelling, G. L. and Coles, C.M., 1996. <u>Fixing Broken Windows: Restoring Order and Crime Our Communities</u>. New York: Free Press.

Kelling, G. L. and Sousa, W.H. Jr., 2001. Do Police Matter? An Analysis of the Impact of New York City's Police Reforms. Manhattan Institute Center for Civic Innovation Civic Report No. 22.

Luria, G., Zohar, D. and Erev, I., 2008. The effect of visibility on effectiveness of supervisory-based interventions. Journal of Safety Research, 39: 273-280.

McCabe, D.L. and Trevino, L.K., 1997. Individual and contextual influences on academic dishonesty: A multicampus investigation. Research in Higher Education, 38: 379-396.

Nee, V., 1998. Norms and networks in economic and organizational performance. American Economic Review, 88: 85-89.

Nevin, J. A. and Grace, R. C., 2000. Behavioral momentum and the law of effect. Behavioral and Brain Sciences, 23: 73–130.

Perry, O, Erev, I. and Haruvy. E., 2002. Frequent probabilistic punishment in law enforcement. Economics of Governance, 3: 71-86.

Rashbaum, W. K., 2002. In new focus on quality of life, city goes after petty criminals. New York Times, May 22, p. 1.

Schneider, A., 1999. Why professors don't do more to stop students who cheat." The Chronicle of Higher Education, 22 Jan.

Silverman, E.B., 1999. New York Police Department Battles Crime. Boston: Northeastern University Press.

Skogan, W., 1990. Disorder and Decline: Crime and the Spiral of Decay in American Cities, New York: Free Press.

Taylor, R., 2001. Breaking away from broken windows: Baltimore neighborhoods and the nationwide fight against crime fear, and decline. Boulder, CO: Westview Press.

Yamagishi, T., 1986. The provision of a sanctioning system as a public good. Journal of Personality and Social Psychology, 51: 110-116.

Zohar, D., 1980. Promoting the use of personal protective equipment by behavior modification techniques. Journal of Safety Research, 12: 78-85.

Zohar, D. and Luria, G., 2004. Climate as a Social-cognitive Construction of Supervisory Safety Practices: Scripts as Proxy of Behavior Patterns, Journal of Applied Psychology. 89: 322-333