Econ. Gov. (2002) 3: 71–86

Economics of Governance
© Springer-Verlag 2002

Frequent probabilistic punishment in law enforcement

Orit Perry[1], **Ido Erev**[2], **and Ernan Haruvy**[3]

[1] Technion, Faculty of Industrial Engineering and Management, Haifa 32000, Israel
[2] Columbia Business School and Technion, Columbia University, Uris Hall, 3022 Broadway, New York, NY 10027, USA
[3] University of Texas at Dallas, 2601 North Floyd Road, Richardson, TX 75083, USA

Received: May 2000 / Acccepted: December 2000

Abstract. Timing and frequency of punishment are critical elements in law enforcement. Previous studies suggest the superiority of immediate punishment schemes over delayed punishment, as well as the importance of frequent punishment. Yet law enforcement schemes which utilize both frequent and immediate punishment are often cost prohibitive. In this work, we propose the "bad lottery immediate punishment" as an effective substitute to immediate punishment. This is a punishment mechanism that signals immediately to an offender that his violation has been spotted, but the actual penalty is delayed and probabilistic. We discuss implications in law enforcement, where probabilistic punishment is potentially more cost effective.

Key words: law enforcement; reinforcement learning

JEL classification: C91, D78, K42

1. Introduction

Experimental studies on the effect of punishment on behavior generally indicate that immediate punishment is far more effective than delayed punishment (e.g., Banks and Vogel-Sprott, 1965; Cohen, 1968; Kamin, 1959; Walters, 1964). Finding applications to this insight in law enforcement, however, is not a straightforward task. Since the optimal delay between the act and the punishment is few seconds, the required implementation tends to be very expensive and, in many cases, impractical. The main goal of the current research is to explore the value

This research was supported by a grant from the USA-Israel Binational Science Foundation and by NASA Cooperative Agreement # NCC 2-995 with Roger Romington and has benefited from insightful conversations with Roger Romington. The authors are indebted to two anonymous referees for helpful comments and advice. All errors and omissions are our sole responsibility.

of potential solutions to this application problem. Two alternative solutions are considered. The first, referred to as *rare immediate punishment* (RIP), is the solution typically selected by law enforcement authorities and planners. For example, traffic police attempt to give immediate punishment (tickets), but only a small portion of the violations is punished. The second solution, referred to as the *Bad Lottery Immediate Punishment*, (BLIP), has become feasible in recent years. Under this solution, the violator receives an immediate signal indicating there is some chance that he or she will be punished. For example, a camera flash that stands for a probability of getting a red light ticket could conceivably function effectively as a cheap and immediate punishment, even though it is merely a (bad) lottery for a distant payoff[1].

The main result of this work is that BLIP is much more effective than RIP with the same expected values. Indeed, BLIP may be as effective as frequent immediate punishment. This result can be predicted from recent studies of the effect of economic incentives on choice behavior, and is supported here by three experiments that examine it directly.

The paper is organized as follows: Section 2 reviews the literature that explores the value of immediate punishment and immediate reinforcements in general. Section 3 summarizes recent findings of regularities that suggest that BLIP is likely to be more effective than RIP. Section 4 presents three experiments intended to evaluate the robustness of this suggestion. Section 5 presents and discusses the experimental results as they relate to the regularities described in Sect. 3. Section 6 concludes.

2. The importance of timing

To illustrate the importance of the timing of punishment on human learning, it is instructive to consider the common problem of remembering to turn off your car headlights following parking. It turns out that the installation of a "light-on" buzzer (producing an annoying sound when the driver turns off the engine before turning the headlights off) reduces the likelihood of forgetting the headlights. That is, the much cheaper immediate punishment (the buzzer) is more effective than the larger delayed punishment received without the installation of the buzzer (the dead battery)[2].

Many laboratory experiments have shown suppression in learning when punishment was delayed. An amusing example is a study by Walters and Demkow (1963) conducted with kindergarten children in Toronto. The experimenter would forbid a child to play with a popular toy and then leave the child alone in the

[1] Red light cameras usually serve as a cost effective deterrent and are not an implementation of the ideas described here. In fact, seeing the flash of a red light camera usually translates to an almost certain punishment, which goes contrary to what is being described in this paper. Nevertheless, the technology does allow for the approach described and is therefore repeatedly mentioned throughout the paper.

[2] It could be argued that the annoying buzzer is simply effective as a reminder and not a reinforcer. In that case, however, one would expect that a pleasant sound would be just as effective.

room and instruct him or her to look through a book while waiting for the experimenter to return. The kids were typically not excited about the book, possibly because it was thick, contained no pictures, and was in the Russian language (Russian is not commonly spoken in Toronto). A loud unpleasant sound from a buzzer followed any attempt to play with the toy. It was shown that when the sound followed immediately it was more effective in discouraging such attempts than the same unpleasant sound with some delay. Other studies showing the importance of immediate punishment are Mowrer and Ullman (1945), Banks and Vogel-Sprott (1965), Cohen (1968), Kamin (1959), and Walters (1964). Applied field research is consistent with these discoveries. For instance, Abramowitz and O'Leary (1990) found that in real classroom settings, children were likely to discontinue misbehavior when a clear verbal reprimand by the teacher was immediate, but were less likely to discontinue misbehavior when the same verbal reprimand was delayed by two minutes. Misbehavior in this setting was defined as verbal or physical interaction with others, standing up, staring into space, or making noise.

3. The economic considerations and the effectiveness of BLIP

Becker (1968) proposed that the relevant criteria in determining optimal reinforcement levels are (1) the size of the punishment, (2) the cost of catching rule violators and (3) the cost of convicting rule violators. Ehrlich (1974) found a "crowding out effect" in law enforcement, whereby an increase in the number of criminal cases handled decreases the effectiveness of the law enforcement agencies due to overload. This finding, together with the cost considerations of Becker, makes the notion of frequent immediate punishment seem impractical at best.

Under the assumption that the timing of the punishment is crucial, the possible solutions for this problem are based on alternative methods to reduce the enforcement cost. This can be done by using less frequent punishment (the "rare immediate punishment," RIP, method), or by using less costly punishment. The less costly punishment considered here is warning signals that stand for a chance of receiving a large punishment later. The administration of this "bad lottery immediate punishment" (BLIP) is particularly easy when automatic equipment is used to identify rule violation. Indeed, in some cases BLIP could be a byproduct of the automated identification process.

As an example, consider the workings of a camera system used to identify drivers who fail to stop at a red light. A red light camera system, strategically positioned at a potential trouble spot, is connected to the traffic-signal system and to sensors buried in the pavement at the crosswalk or stop line. The camera system continuously monitors the traffic signal and is triggered when any vehicle passes over the sensors faster than a preset minimum speed and at a specified elapsed time after the signal has turned red. A second photograph is taken which shows the violator in the intersection. The camera records the date, time of day, time elapsed since the beginning of the red signal, and the speed of the vehicle. As a byproduct of the photographing process the offender is notified of being spotted by a bright flash emanating from the camera.[3]

Three behavioral regularities, discovered in experimental economic research, suggest that BLIP is likely to be more effective than RIP. The first is the finding that *unrealized gambles can serve as reinforcements*, and their effect is very similar to the effect of monetary reinforcements with the same expected value. One indication for this regularity comes from experimental studies paying subjects in lottery tickets, or probability points, instead of in real money. Smith (1961) was first to discuss this equivalence in a theoretical framework. Roth and Malouf (1979) were first to use such payment schemes in experiments in order to induce risk neutrality. Since then, studies have shown that binary lotteries induce behavior equivalent to that resulting from direct payment (e.g., Cox, Smith, and Walker, 1985, 1988; Walker, Smith, and Cox, 1990; Cox and Oaxaca, 1995).[4] In the negative payoff domain, there are far fewer studies. However, one such study is highly relevant to the study at hand: Barkan (1998) examined the effect of a red light that signaled a bad gamble and found that this effect was nonnegligable.

The second regularity involves the effect of *observed payoff variability*. Previous studies (Thaler et. al., 1997; Erev et. al., 1999; Haruvy and Erev, 2000; Barron and Erev, 2000) show that observed payoff variability can slow down learning. Thus, since RIP increases observed payoff variability[5] (relative to immediate punishment with the same expected value), it is expected to slow down the learning process (the response to the punishment).

A third finding is related to the effect of recent outcomes on future choices. Experimental results show a *tendency to choose the best reply to the most recent outcomes* (e.g., Barron & Erev, 2000). This tendency implies an underweighting of rare outcomes and, for that reason, is expected to impair the effectiveness of RIP. Another indication for the tendency to ignore rare outcomes comes from the study of policy and legal issues. For example, Block and Gerety (1995) have found that criminals are more sensitive to the frequency and probability of punishment than they are to the size of punishment.

Given the first regularity, it is natural to assume that lotteries with negative expected values ("bad lotteries") can serve as immediate punishment. It seems

[3] Although these programs are mainly used for cost-effectiveness, they are nonetheless effective in reducing offense rates. Several U.S. red light camera enforcement programs reported a decline in the number of tickets issued over time, suggesting these programs are effective in reducing red light violation rates (e.g., Retting et al, 1999a, 1999b).

[4] Of course, in this type of experiments, different choices result in x probability points versus y probability points. Hence arguments of risk aversion are not relevant. However, in the BLIP setting, the payments are in terms of both lotteries (in BLIP) and certain payoffs. Here, risk aversion certainly enters, although we are quick to assume it away. These examples do, however, illustrate that lotteries can serve as effective reinforcements.

[5] Perhaps the emphasis should be on the word *observed*. If the decision is finitely repeated, and the probability of punishment is extremely small, so that punishment is not likely to be observed by the vast majority of violators, observed payoff variability under RIP would be near the zero punishment payoff variability for most participants. Hence, the effect of RIP on payoff variability depends on the exact probability of punishment.

that a bad lottery, such as that signaled by a camera flash at a red light intersection, may possess the power of punishment without incurring the cost of congestion on law enforcement agencies in arresting, trying, convicting, and administering punishment to a large volume of felons. This suggestion implies that the effect of BLIP is expected to be similar to the effect of immediate frequent punishment (FIP) with the same expected value.

Is BLIP more effective than the use of rare immediate punishment (RIP)? That of course depends on what we define as effective. One definition of 'effective,' the one of relevance to psychologists, would pick the system that induces faster learning towards the best response. The other definition, of greater relevance to law enforcement agencies, would pick the system that induces a lower rate of unlawful behavior over time.

The current analysis takes the point of view of law enforcement agencies. To determine whether BLIP is more effective than RIP, we must consider the two effects of variability and underweighting of rare outcomes. As noted above, under the assumption of the same expected value of each choice across systems, the underweighting of rare outcomes is expected to impair the effectiveness of the RIP system, since potential violators' perceived probability of rare punishment is lower than the true probability. The effect of observed payoff variability is less evident. A high variability in observed payoffs is expected to reduce learning speed toward expected value maximization. Thus, to the extent that BLIP reduces variability relative to RIP, it is expected to facilitate enforcement if the expected value from the unlawful choice is lower than the expected value from the lawful choice. When the expected value of the punishment is not high enough to make the lawful choice higher in expected value, the reduced variability effect of BLIP (and FIP) could impair law enforcement effectiveness by making criminals learn faster that crime pays.

These observations imply that BLIP is likely to be the optimal method when the punishment is high enough (in which case the two effects work in the same direction). However, careful quantification of the relative value of the two effects of variability and rarity is required to identify the optimal method when violating the law maximizes expected value (in which case the two effects contradict each other). To illustrate this point it is constructive to consider numerical examples:

Example 1. Consider a driver expected to reach her target in 1 to 2 minutes if she runs the next traffic light, and in 2 to 3 minutes if she stops at the light. Example 1 (top left of Figure 1) presents the driver's decision problem under the assumption that the fine for running the light can be translated (from the viewpoint of the driver) to the cost of 200 minutes and the probability of a fine is 1/100. In addition, it is assumed that BLIP is implemented by identifying one-third of the violators and signaling to them that they acquired the bad lottery $\{-200 \text{ with } p = 0.03 \text{ and } 0 \text{ with } p = 0.97\}$. Notice that in this example, the expected value of the unlawful choice (-3.5) is lower than the expected value of the lawful behavior (-2.5). Thus, the current analysis implies that BLIP is expected to be more effective than RIP. This prediction is presented in the first column, graphs 1 and 4, of Figure 1 under the assumption that drivers learn according to the quantification of reinforcement learning proposed by Barron and Erev (2000).[6] This model, one possible abstraction of the effects presented above, is described in Appendix 1. Graph 1 presents the predictions under the assumption of uniform initial tendencies. In graph 4, the initial propensities are derived from the variant of cumulative prospect theory (Tversky & Kahneman, 1992) proposed by Barron and Erev (2000). A comparison of the two sets of predictions reveals that the predictions are robust to the assumed initials.

Notice that the current predictions are not trivial. Indeed, two of the leading approaches to decision making under uncertainty, expected utility theory with risk aversion and prospect theory, make very different predictions than the ones espoused here. Traditional expected utility maximization with generally accepted estimates of risk aversion would imply no unlawful choices under any of the three systems of punishment. Moreover, the difference between the expected utilities of lawful and unlawful choices is the smallest under FIP. Hence, an expected utility model with a noisy payoff-sensitive best-response mapping would predict FIP to be the least effective in combating crime. Cumulative Prospect Theory (Tversky and Kahneman, 1992) implies *risk seeking* in the loss domain, but it also predicts *overweighting of rare events*.[7] With the parameters that best fit the median subject in Tversky and Kahneman (1992), the latter effect is more pronounced. Thus, a strict generalization of this theory to the current task implies that RIP should be the most effective method. Hence, these two very different and virtually opposing approaches both predict RIP to be superior to FIP and BLIP.

Another nontrivial aspect of the current set of predictions is the fact that the expected advantage of BLIP over RIP, when the expected value of the lawful action is highest, is not a result of the warning signals by themselves (i.e., the red lights), but rather of the information conveyed by these signals. Indeed, the advantage of BLIP over RIP remains when the same warning signals are presented under both. Learning research indicates that the warning signals will have a major impact under BLIP, where they correspond to future punishment possibilities and hence serve as reinforcement. Yet they will be ignored by subjects in the RIP treatment, since they do not correspond to future punishment there.

Example 2. Consider a variant of Example 1 in which the size of the fine is reduced to the equivalent of 50 minutes (second column of Figure 1). This change increases the expected value of breaking the law to -2. As a result, the two effects (payoff variability and underweighting of rare events) point to contradicting directions. Figure 1, second-column, graphs 2 and 5, shows that with the parameters estimated by Barron and Erev (2000), the model predicts a

[6] In the simulations, the effect of the bad lottery is assumed to equal its expected value. To facilitate experimental investigation of these examples we replaced the uniform arrival time distributions [2,3] and [3,4] with the discrete gambles (2 with $p = 0.5$; 3 with $p = 0.5$) and (3 with $p = 0.5$; 4 with $p = 0.5$), respectively.

[7] Overweighting of rare events is not assumed by Barron and Erev's initial propensities variant of prospect theory.

	Experiment 1			Experiment 2			Experiment 3	
	Unlawful	Lawful		Unlawful	Lawful		Unlawful	Lawful
RIP	−1 (0.5), −2 (0.5) plus –200 (0.01)	−2 (0.5), −3 (0.5)	RIP	−1 (0.5), −2 (0.5) plus –50 (0.01)	−2 (0.5), −3 (0.5)	RIP	−1 (0.5), −2 (0.5) plus –200 (0.01)	−4 (0.5), −5 (0.5)
FIP	−1 (0.5), −2 (0.5) plus –6 (0.33)	−2 (0.5), −3 (0.5)	FIP	−1 (0.5), −2 (0.5) plus –1.5 (0.033)	−2 (0.5), −3 (0.5)	FIP	−1 (0.5), −2 (0.5) plus –6 (0.033)	−4 (0.5), −5 (0.5)

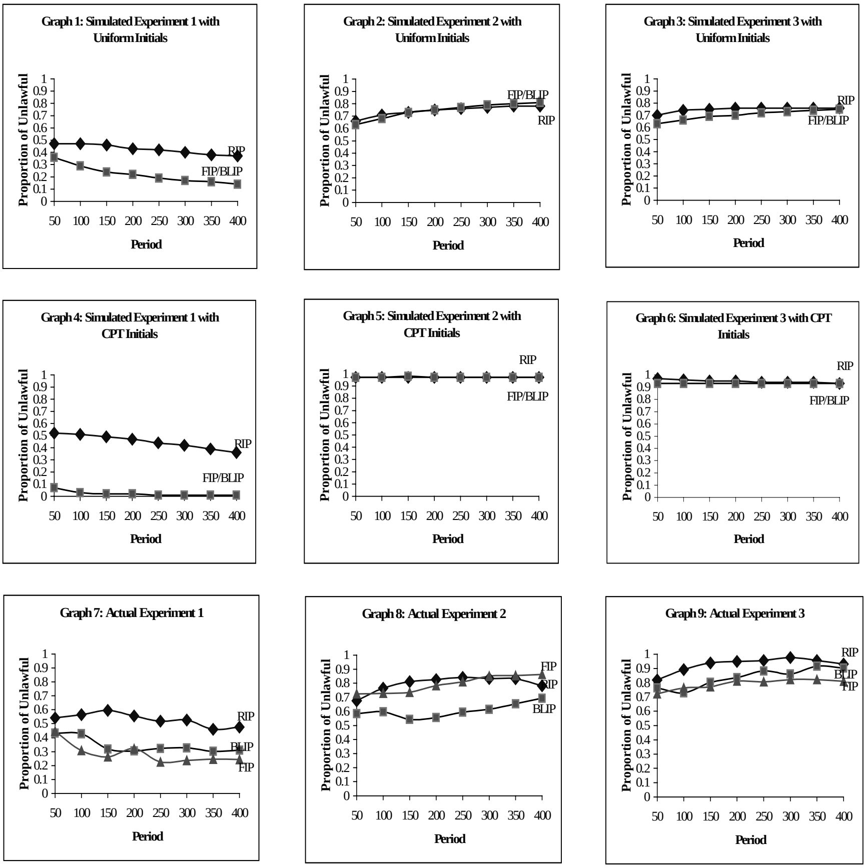

Fig. 1. The experiments and simulations

very small difference between RIP and BLIP. Moreover, the effect changes with time and shows some sensitivity to the assumed initials.

The next section presents three experiments that test the robust predictions of the current analysis and collect more data to improve our understanding of the relative importance of the two effects described above.

4. Experimental design

Nine experimental treatments were run. Each treatment involved subjects facing the same two choices repeatedly for 400 periods. Choices were represented by two virtual buttons on a computer screen. The buttons were labeled "Right" and "Left" and subjects were told to select a button in each repetition. Once pressed, the button selected would display the token payoff for the corresponding choice. The token payoff was always negative due to the nature of the investigation. However, subjects all started out with a positive endowment. The number of tokens lost in a given round would be subtracted from the cumulative payoffs displayed at the bottom of the screen. Although the buttons were presented to the subjects in neutral terms, we find it convenient, for the purpose of continuity of the discussion from the last section, to refer to the choices as 'the undesirable behavior' (such as running a red light) and 'the desirable behavior.' Similarly, though payoffs were in tokens (conversion rate of 40 tokens = 1 Shekel = $0.25), we may think of payoffs as representing time lost in traffic, or time-equivalent punishment.

For exposition purposes we classify the nine *treatments* into three *experiments*, involving three *conditions* each. The three conditions in a given experiment have the same expected payoff for a given choice, but differ in the frequency and immediacy of punishment. The three conditions correspond to the RIP, BLIP, and FIP mechanisms described earlier. Specifically, in the RIP condition, the final payoff for a given trial is provided immediately following the choice. In addition, subjects see a red light one-third of the times they select the unlawful action. The BLIP condition is identical to RIP with the exception that the immediate feedback does not entail sure monetary punishment. Rather, subjects are informed that each of the warnings that appear on the screen stands for a 0.03 chance of getting the large punishment at the end of experiment.[8] In the third condition, FIP, subjects receive an immediate payoff set equal to the expected value of the warning in the BLIP condition (which is 6 tokens in the first and third experiments and 1.5 tokens in the second experiment).

The three experiments differed in the expected payoff for the two choices as well as in the initial endowment. In the first experiment, corresponding to our first example in Sect. 2, the lawful choice had the higher expected value relative to the unlawful choice (-2.5 versus –3.5). In the other two experiments, the unlawful choice had the higher expected value. Payoff tables for treatments RIP and FIP in the three experiments are presented on the left-hand side of Figure 1. The payoff structure of the BLIP condition is identical to the payoff structure of the RIP condition, with the exception that the bad gambles are delayed. Initial endowments are 2000 tokens in the first two experiments and 2400 in the third experiment.

In each of the first two experiments, 25 subjects participated in each treatment. In the third experiment, 18 subjects were assigned to each treatment. No subject

[8] This is the probability of getting punished conditional on having seen the warning. Since the warning occurs with probability of 1/3, the unconditional punishment probability is 0.01.

participated in more than one treatment. A total of 204 subjects participated in this study. Subjects were all Technion students.

5. Results

The graphs in the last row of Figure 1 display the patterns of adjustment in terms of the proportion of unlawful behavior, in each treatment and each experiment. Like the top two rows that present the corresponding simulations of the model, they represent the average proportion of unlawful choices in eight blocks of 50 periods over (the 25) participants in each treatment[9]. Table 1 reports the proportion of unlawful choices over all players and all periods in each treatment of each experiment.

Table 1. Proportion of "unlawful" choices over all players and all periods in each treatment of each experiment

	Treatment	**Mean**	**Standard Deviation**	**Significance Tests t-statistic & P-value (2-tail)**	
				RIP vs. BLIP	**BLIP vs. FIP**
Experiment 1	**RIP**	0.529	0.291	$t[48] = 2.128$ $p = 0.038$	
	BLIP	0.343	0.325		$t[48] = 1.393$ $p = 0.170$
	FIP	0.286	0.195		
Experiment 2	**RIP**	0.795	0.200	$t[48] = 2.350$ $p = 0.022$	
	BLIP	0.604	0.352		$t[48] = 2.297$ $p = 0.026$
	FIP	0.793	0.214		
Experiment 3	**RIP**	0.927	0.084	$t[34] = -1.919$ $p = 0.061$	
	BLIP	0.836	0.179		$t[34] = -1.116$ $p = 0.270$
	FIP	0.790	0.268		

We first note the surprising lack of significance in the difference between BLIP and FIP. Over the three different experiments presented in this work, the difference between BLIP and FIP is insignificant (t [132]= −0.78, p-value of 0.4362). As we will see in the paragraph that follows, this stands in sharp contrast to the magnitude and statistical significance of the difference between BLIP and RIP. Recall that the present model considered lottery tickets to be equivalent to reinforcement of an equal expected value. While the results are not necessarily an indication that this assumption is correct, they do suggest that this assumption provides us with ability, in the lack of a better alternative, to compare a lottery

[9] The eight blocks are for exposition purposes only, to avoid cluttering the graphs with visually cumbersome information. During the experiment, subjects played 400 periods without interruption.

system to a system of immediate punishment. This has major implications in real-life settings, where punishment is too infrequent to be effective. Though it is clear that many smaller negative reinforcements are better than an infrequent large one of the same magnitude, often a large punishment is impractical to administer in small portions, or possibly is indivisible. The indivisibility of a positive outcome has often been cited as a reason for why people buy lottery tickets in the positive domain (Friedman and Savage, 1948; Ng, 1965; Robson, 1996). For example, a house, a car, and college education are large indivisibles. On the negative reinforcement side, punishments such as eviction of a tenant, revocation of a driver's license, or expulsion from school, are likewise indivisible. The possibility of lotteries in the negative domain raises a possible solution to the frequency of punishment problem wherein large rare punishments are stretched in effectiveness through the use of lotteries.

Whereas BLIP and FIP are not statistically different, the same cannot be said for BLIP and RIP. The results show that for all three experiments the proportion of lawful choices is higher under the BLIP treatment than under the RIP treatment. This difference between RIP and BLIP is highly significant over the three experiments ($t[132] = -3.58$, p-value of 0.0005).

In the first experiment, the model's prediction (as seen in Figure 1 to the right) is unambiguously in favor of BLIP. Over the 400 trials, participant's average proportion of unlawful choices was 0.343 (Stdev = 0.325) in the BLIP treatment and 0.529 (Stdev = 0.291) in the RIP treatment. The difference is significant ($t[48] = 2.128$, p-value of 0.038). Hence, the experimental results support the prediction of the model.

In the second example (experiment), as discussed in Sect. 2, the model provides mixed predictions. If our goal were to slow down learning toward unlawful behavior (the best response action in this scenario), we would wish to generate high variability in payoffs, which RIP would provide. However, due to the rarity effect that comes with RIP, causing people over time to discount punishment, BLIP may nonetheless be preferred. The simulations show relatively small differences between the conditions. The actual data, on the other hand, shows BLIP to lead over RIP in proportion of lawful behavior, in contrast to the model's prediction. In that sense, the model has clearly overestimated the effect of variability relative to rarity. Participants' average proportion of unlawful choices was 0.604 (STD = 0.352) in the BLIP treatment and 0.795 (STD = 0.2), in the RIP treatment. The difference is significant ($t[48] = 2.35$, p-value of 0.022).

The third experiment is very similar to the second one, as the expected value of the unlawful behavior is greater than that of the lawful one. The difference lies in the degree of variability, which is much higher in the third experiment relative to the second. Like in the second experiment, the simulations predict BLIP and RIP to be virtually indistinguishable. Participants' actual average proportion of unlawful choices was 0.836 (STD = 0.179) in the BLIP treatment and 0.927 (STD = 0.084) in the RIP treatment. This difference is not significant at the 5% level but is significant at 10% ($t[34] = -1.934$, p-value of 0.061).

6. Conclusions

The analysis of the optimal punishment in law enforcement has led psychologists on the one hand and economists on the other to provide valuable insights. Psychology research suggests that the immediacy of punishment is the crucial variable and should be emphasized, whereas economic analysis suggests that the cost– to the enforcer as well as to the violators– is the critical variable. Unfortunately, these recommendations appear to be incompatible. The current research does not challenge these recommendations. Rather, it asks how they can be jointly implemented. Building on recent research in experimental economics, we suggest that the task can be simpler than it seems. Since (1) bad unrealized lotteries can serve as punishments, and (2) decision makers tend to underweight rare events, the optimal solution calls for using frequent warning signals (that correspond to delayed punishments) as immediate punishments. Three experiments validate the plausibility of this solution.

There is some indication that rule enforcement agencies are capable of adopting the solution supported here. One example of where implementation of the proposals of this work could prove useful is in red light cameras. Recent years have witnessed the increased use of red light camera systems around the U.S. (Retting et al, 1999a, 1999b) and Europe. Yet, it is easy to see that the implementation can be improved. For example, the red light camera near our university is programmed to release the first flash only when a driver enters the junction more than 1.5 seconds following the red light appearance. As a result, very few violators see the warning (and almost all of them end up receiving the ticket). It seems that the local police treats the existence of the camera as the warning signal, and wants to save "flashes" (or retain the high correlation between observed flashes and tickets). The current analysis suggests that the effectiveness of this specific camera can be improved by changing the cutoff for the release of the first flash even without changing the ticketing rule. We contend that drivers who enter the junction in red do not know their exact entering time. Thus, additional flashes will serve as BLIP.

Although the current paper studies a very focused practical problem, it leads to some interesting theoretical suggestions. Most importantly, it shows that learning research can lead to robust predictions that cannot be derived from the assumption of expected utility maximization with loss aversion or from a strict generalization of Prospect Theory. In addition, the results of the three experiments suggest that the tendency to underweight rare events is stronger than the tendency predicted by Barron and Erev's (2000) quantification. Thus, the modeling of human adjustment to economic incentives that facilitates the current analysis can be improved.

References

Abramowitz, A. J., O'Leary, S. G. (1990) Effectiveness of delayed punishment in an applied setting. Behavior Therapy 21: 231–239.

Ammons, R. B. (1956) Effects of knowledge of performance: A survey and tentative theoretical formation. Journal of General Psychology 54: 279–299.

Banks, R. K., Vogel-Sprott, M. D. (1965) The effect of delayed punishment on an immediately rewarded response in humans. Journal of Experimental Psychology 70: 357–359.

Barkan, R. (1998) Decreasing risk-taking: The reinforcement value of non-payoff information in accident prevention tasks. Indiana University working paper.

Becker, G. S. (1968) Crime and punishment: An economic approach. Journal of Political Economy 76(2): 169–217.

Barron, G., Erev, I. (2000) Toward a general descriptive model of one-shot and repeated decision making under risk and uncertainty. Technion working paper.

Block, M. K., Gerety, V. E. (1995) Some experimental evidence on differences between student and prisoner reactions to monetary penalties and risk. The Journal of Legal Studies 24(1): 123–138.

Cohen, P. S. (1968) The interactive effects of delay and intensity of shock. Journal of the Experimental Analysis of Behavior 11, 789–799.

Cox, J.C., Oaxaca, R. L. (1995) Inducing risk-neutral preferences: further analysis of the data. Journal of Risk and Uncertainty 11: 65–79.

Cox, J. C., Smith, V. L., Walker. J. M. (1985) Experimental development of sealed-bid auction theory; calibrating controls for risk aversion. American Economic Review 75(2): 160–165.

Cox, J. C., Smith, V. L., Walker, J. M (1988), Theory and individual behavior of first-price auction. Journal of Risk and Uncertainty 1(1): 61–99.

Ehrlich, I., (1974) Participation in illegitimate activities – An economic analysis. In: Becker, G. S., Landes, W. M. (eds.) The Economics of Crime and Punishment. Columbia University Press, New York: 68–134.

Erev, I., Roth, A. (1998) Predicting how people play games: Reinforcement learning in experimental games with unique mixed strategy equilibria. American Economic Review 88(4): 848–881.

Erev, I., Bereby-Meyer, Y., Roth, A. (1999) The effect of adding a constant to all payoffs: Experimental investigation, and implications for reinforcement learning models. Journal of Economic Behavior and Organization 39(1): 111–128.

Friedman, M., Savage, L. J. (1948) The utility analysis of choice involving risk. Journal of Political Economy 56: 279–304.

Grogger, J. (1991) Certainty versus severity of punishment. Economic Inquiry 29(2): 297–309.

Haruvy, E., Erev, I. (2000) When to pursue variable pricing: The relationship between price format and quality, working paper, Technion.

Ilgen, D. R., Fisher, D. C., Taylor, M. S. (1979) Consequences of individual feedback on behavior in organizations. Journal of Applied Psychology 64: 349–371.

Kahneman, D., Tversky, A. (1979) Prospect theory: An analysis of decisions under certainty. Econometrica 47(2): 263–291.

Kamin, L. J. (1959) The delay of punishment gradient. Journal of Comparative and Physiological Psychology 52: 434–437.

Mowrer, O. H., Ullman, A. D. (1945) Time as a determinant of integrative learning. Psychological Review 52: 61–90.

Ng, Y. K. (1965) Why do people buy lottery tickets? Choice involving risk and the indivisibility of expenditure. Journal of Political Economy 73: 530–535.

Retting, R. A., Williams, A. F., Preusser, D. F., Weinstein, H. B. (1995) Classifying urban crashes for countermeasure development. Accident Analysis and Prevention 27: 283–294.

Retting, R. A., Williams, A. F., Farmer, C. M., Feldman, A. F. (1999a) Evaluation of Red Light Camera Enforcement in Oxnard, California. Accident Analysis and Prevention 31: 169–174.

Retting, R. A., Williams, A. F., Farmer, C. M., Feldman, A. F. (1999b) Evaluation of Red Light Camera Enforcement in Fairfax, Va., USA. ITE Journal, August: 30–34.

Robson, A. (1996) The evolution of attitudes to risk: Lottery tickets and relative wealth. Games and Economic Behavior 14(2): 190–207.

Roth, A. E., Malouf, M. (1979) Game theoretic models and the role of information in bargaining. Psychological Review 86: 574–594.

Smith, A. (1961) Consistency in statistical inference and decision. Journal of the Royal Statistical Society 23:1–25.
Thaler, R., Tversky, A., Kahneman, D., Schwartz, A. (1997) The effect of myopia and loss aversion on risk taking: An experimental test. Quarterly Journal of Economics 112(2): 647–661.
Tversky, A., Kahneman, D. (1992) Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty 5(4): 297–323.
Walker, J. M., Smith, V. L., Cox, J. C. (1990) Inducing risk neutral preferences: An examination in a controlled market environment. Journal of Risk and Uncertainty 3(1): 5–24.
Walters, R. H. (1964) Delay of reinforcement gradients in children's learning. Psychonomic Science 1: 307–308.
Walters, R. H., Demkow, L. (1963) Timing of punishment as a determinant of response inhibition. Child Development 34: 207–214.

Appendix A: Barron and Erev's (2000) model

Barron and Erev's model is an extension of a joint generalization of cumulative prospect theory (Tversky & Kahneman, 1992) and Erev et al.'s (1999) reinforcement learning model. Its basic assumptions are presented below:

Strategies and decisions:
The decision makers are assumed to consider two sets of strategies: Stage game strategies (left button and right button in the current study), and cognitive strategies. The cognitive strategies are rules that condition the action in trial t + 1 on the outcome of the first t trials. Three cognitive strategies are assumed: (1) Best reply to the last period, (2) best reply to some pattern (which is reduced to expected value maximization in the current context), and (3) selection of the action that leads to the lower proportion of losses (implies indifference over actions in the current study).

In each trial, the decision maker is assumed to select one set of strategies, stage game or cognitive, and then to select one of the strategies in the selected set. Typically, the selected strategy implies an act. When a cognitive strategy is selected that implies indifference (in the current setting, the loss aversion strategy implies indifference), one of the stage game strategies is selected at random. Both selection phases are assumed to obey the same choice rule.

Initial propensities:
Under the*assumption of uniform initials*, the initial propensity to select each of the sets and each of the strategies equals A(1) – the expected payoff from random choice over stage game strategies.

The *assumption of prospect theory-like initials* implies that the propensity to use strategy j is: $q_j(1) = SV_j$ + D, where SV_j is the subjective value of the action implied by strategy j (according to Cumulative Prospect Theory, where strategy j can be either a cognitive strategy or a stage game strategy). To retain the mean of the initial propensities (over both cognitive and stage game strategies) at $A(1)$, we let $D = A(1) - \left(\frac{1}{5}\right) \sum_{j=1}^{5} SV_j$.

The initial propensity to use each of the sets of strategies is equal to the mean propensity of the strategies in that set. We let δ denote the set.

Average updating:

The propensity to select strategy set δ in round $t+1$ is a weighted average of the initial propensity of set $\delta(q_\delta(1))$ and the average payoff obtained from playing set δ in the first t rounds ($AVE_\delta(t)$). The weight of the initial propensity is a function of a "strength of initial propensities" parameter $N(1)$. The weight of the average past payoff is a function of the number of times set δ was actually chosen in the past ($C_\delta(t)$). Specifically,

$$q_\delta(t+1) = q_\delta(1)\frac{N(1)}{C_\delta(t)+N(1)} + AVE_\delta(t)\frac{C_\delta(t)}{C_\delta(t)+N(1)} \tag{1}$$

The propensity of the selected strategies is updated by the same rule with N(1) replaced by N(1)/m_δ where m_δ is the number of strategies in the set.

Exponential Response Rule:
The probability $p_\delta(t)$ that a decision maker selects set δ at time t is given by,

$$p_\delta = e^{q_\delta(t)\lambda/S(t)} \Big/ \sum_{k=1}^{2}\left(e^{qk(t)\lambda/S(t)}\right) \tag{2}$$

where the sum is over the two sets, λ is a parameter that determines reinforcement sensitivity, and $S(t)$ is a measure of the standard deviation of the payoffs that the decision maker has experienced up to time t. The selection among strategies *within each set* obeys the same rule.

Thus, the probability of selecting a given set and strategy increases with the propensity to select it (which increases with the average payoff from past selections). The division by the deviation measure, $S(t)$, implies that noisy reinforcements reduce reinforcement sensitivity (and thus lead toward more uniform choice probabilities).

The deviation, $S(t)$, is estimated by the average absolute difference between the recent payoff (x at trial t) and the accumulated average payoff at the first t trials ($A(t)$). Following the logic of Equation 1, $S(t)$ is updated as follows:

$$\mathrm{S(t+1) = S(t)W'(t) + |A(t) - x|_{(1} - W'(t))} \tag{3}$$

where $W'(t) = (t + 2N(1))/(t + 2N(1) + 1)$. For the deviation's initial value, $S(1)$, we take the expected absolute difference between the payoff from random choice and "the expected payoff given random choice."

The average payoff $A(t)$ is calculated in a similar manner:

$$A(t+1) = A(t)W'(t) + x(1 - W'(t)) \tag{4}$$

where $A(1)$ is the expected payoff from random choice.

Altogether, the model has two learning parameters. Barron and Erev's estimation yields the values: $\lambda = 4.2$ and $N(1) = 100$. The assumption of prospect theory-like initials implies three additional parameters. Barron and Erev's estimation yields the values: loss aversion (λ in Tversky and Kahneman, μ in Barron and Erev) = 2.25, effective weighting function (γ) = 1, diminishing initial sensitivity (α) = 0.78.

Appendix B. Instructions for Experiment 1 (Translation from Hebrew)

Instructions for the RIP condition

Welcome and thank you for your participation. You are about to participate in an experiment of 400 rounds. In each round you must choose between the two virtual buttons on your screen, one labeled “Right” and the other labeled “Left.” To select one of the virtual buttons, you must move the mouse cursor to a position on top of your selected virtual button and click the left mouse button. Following each choice, you will receive an output in the upper window on the screen, specifying the number of tokens you lost in that round. You will also receive output in the bottom window on your screen specifying your total cumulative token earnings so far. Following the output for the first round, you must make another choice, await the output, and so on, until the end of the experiment.

For each choice of the right button, you have a 50% chance of losing 2 tokens and 50% chance of losing 3 tokens. For each choice of the left button, you have a 50% chance of losing 1 token and 50% of losing 2 tokens. *In addition, one third of the times you press the left button, you will see the button turning red. When that happens, you have a 3% chance of losing 200 tokens in addition to the 1 or 2 tokens already lost. Should the loss of 200 tokens materialize, you will be notified immediately and the loss will be immediately subtracted from your cumulative earnings.*

You begin with 2000 tokens. The value of each token is 0.025 NIS.

Instructions for the BLIP condition

Welcome and thank you for your participation. You are about to participate in an experiment of 400 rounds. In each round you must choose between the two virtual buttons on your screen, one labeled “Right” and the other labeled “left.” To select one of the virtual buttons, you must move the mouse cursor to a position on top of your selected virtual button and click the left mouse button. Following each choice, you will receive an output in the upper window on the screen, specifying the number of tokens you lost in that round. You will also receive output in the bottom window on your screen specifying your total cumulative token earnings so far. Following the output for the first round, you must make another choice, await the output, and so on, until the end of the experiment.

For each choice of the right button, you have a 50% chance of losing 2 tokens and 50% chance of losing 3 tokens. For each choice of the left button, you have a 50% chance of losing 1 token and 50% of losing 2 tokens. *In addition, one third of the times you press the left button, you will see the button turning red. When that happens, you have a 3% chance of losing 200 tokens in addition to the 1 or 2 tokens already lost. Should the loss of 200 tokens materialize, you will* **NOT** *be notified until the end of the experiment and the loss will only at the end of the experiment be subtracted from your cumulative earnings.*

You begin with 2000 tokens. The value of each token is 0.025 NIS.

Instructions for the FIP condition

Welcome and thank you for your participation. You are about to participate in an experiment of 400 rounds. In each round you must choose between the two virtual buttons on your screen, one labeled "Right" and the other labeled "left." To select one of the virtual buttons, you must move the mouse cursor to a position on top of your selected virtual button and click the left mouse button. Following each choice, you will receive an output in the upper window on the screen, specifying the number of tokens you lost in that round. You will also receive output in the bottom window on your screen specifying your total cumulative token earnings so far. Following the output for the first round, you must make another choice, await the output, and so on, until the end of the experiment.

For each choice of the right button, you have a 50% chance of losing 2 tokens and 50% chance of losing 3 tokens. For each choice of the left button, you have a 50% chance of losing 1 token and 50% of losing 2 tokens. *In addition, one third of the times you press the left button, you will see the button turning red. When that happens, you will lose 6 tokens in addition to the 1 or 2 tokens already lost.*

You begin with 2000 tokens. The value of each token is 0.025 NIS.