Applied Game Theory. 1979 ©Physica-Verlag, Wuerzburg/Germany

An Optimal Conviction Policy for Offenses that May Have Been Committed by Accident

By *A. Rubinstein*, Jerusalem[1])

Abstract: When the tax authorities discover that a taxpayer has failed to report a certain part of his income, they cannot tell whether this is the result of deliberate tax evasion or, perhaps, the result of an innocent oversight. Penalties must be designed so as to apply whenever misreporting is discovered, but society may very well wish to distinguish between a deliberate offense and an offense that has been committed by accident and to be more lenient in the latter case. However, leniency will encourage people to commit the offense deliberately. In a one-shot game, equilibrium consists of society picking severe penalties and innocent offenders being hit hard. However, in the repeated game, there exists an equilibrium in which the optimal penalties imposed by society on people with a "reasonable" record are lenient and the optimal strategy for the individuals is to refrain from deliberate offenses.

1. Introduction

In almost all criminal proceedings, some doubt remains as to the guilt of the accused. Often the factual element (actus reus) is not completely conclusive and even if it is, the mental element necessary in order for an act to become a crime in law (mens rea) is also frequently questionable.

Examples are numerous; here is a short list:

1. The windscreen of a car in a parking lot carries a parking permit with yesterday's date. The owner claims it has been backdated by mistake.
2. A source of income which does not appear in the appropriate income tax return is discovered. The declarer swears that he had omitted it out of forgetfulness, despite his efforts to give a true account of his earnings.
3. One of the headlights of a car is unlit. Conceivably it had just failed, and the driver who had checked the headlights before driving away, had found it to be in working order.
4. A person is caught leaving a supermarket with unpaidfor merchandise in his possession. Two doubts may arise; perhaps the person was confused or possibly someone else might have placed the merchandise where it was found without the person's knowledge.

In these cases, a complete dispelling of the doubt is almost impossible and the legal system is faced with a dilemma; conviction may risk injustice whereas acquittal may open the door to widespread breach of the law.

---

[1]) *Ariel Rubinstein*, Department of Economics, The Hebrew University of Jerusalem, Israel.
My thanks to Professor M.E. Yaari for his valuable comments and encouragement.

As far as once-and-for-all situations are concerned (as for example in cases of crimes like murder) the law has no alternative but to consider which is more unpalatable: A possible miscarriage of justive or a weakening of the deterrent effect of the penalty. However, for offenses which an individual might commit periodically, a wide range of conviction and punishment policies, built around the offender's record, are possible.

In this paper, I shall consider a simple situation which has at its core the social dilemma described above. It will be described as a game in section 2. In section 3 I shall discuss the corresponding repeated game in order to examine the possible policies which the legal system may adopt. The main result (presented in section 4) is that in the repeated game, there exists a pair of strategies, one for the penal system and one for the individual, which are jointly optimal and which have the following structure: In any period, if the individual is discovered as having committed the offense, then he is penalized only if his long-run record is "unreasonably" bad; as for the individual, his optimal strategy is to refrain from committing the offense deliberately.

2. The Isolated Game

Consider the following two-person game whose players are society (player 1) and a typical member of society (player 2).

The individual has to choose between two kinds of behaviour:

B – Committing a given act which is advantageous to himself but harmful to society at large.

G – Refraining from this act.

Even if the individual chooses G, the act may still be committed, by a cause outside his control. Let us think of this unintentional committing of the act as the result of a move by a fictitious *chance player* who picks the move "+" ("commit") with a given probability, α, and the move "–" ("do not commit") with probability $1-\alpha$, where $0 < \alpha < 1$. The chance player's set of strategies, that is the set, $\{+, -\}$ will be denoted S_c.

Society has complete information about whether an act has or has not been committed, but it has no way of telling whether the act was committed wilfully or accidentally. Suppose that there are but two possible sentences which the court can pass when the act has been committed, namely conviction with a fixed penalty, or acquittal.

Society's set of strategies – S_1 – contains two elements:

P – Convict and punish the individual if the act has been committed.

NP – Do not convict the individual even though the act has been committed.

We assume the rules of the legal system are a matter of public knowledge, i.e., player 1 announces his strategy at the beginning of the game. The individual's strategy is therefore of the form: "I will do X if society chooses P, and Y otherwise", where $X, Y \in \{B, G\}$. Let this strategy be denoted simply XY. Then, the individual's set of strategies – S_2 – contains four elements: GG, GB, BG and BB.

The following outcomes correspond to the various choices that the players and the chance player can make:

Outcome 1: $(NP, G, -)$. Society is lenient and the act is not committed.
Outcome 2: $(P, G, -)$. Society is strict and the act is not committed.
Outcome 3: $(NP, G, +)$. The act is committed unintentionally and the offender is not punished.
Outcome 4: $(P, G, +)$. The act is committed unintentionally but the offender is punished.
Outcome 5: (NP, B). The act is committed deliberately, but the individual is not punished.
Outcome 6: (P, B). The act is committed deliberately, and the individual is punished.

Let the set consisting of these six outcomes be denoted S, and let u_1 and u_2 be respectively the utility functions of society and of the individual, with $u_i: S \to \mathbf{R}$ $i = 1, 2$.
Assume that u_1 and u_2 are given by:

$$u_1\ (NP, G, -) = u_1\ (P, G, -) = 4$$
$$u_1\ (NP, G, +) = 3$$
$$u_1\ (P, B) = a \text{ with } 1 < a < 3$$
$$u_1\ (NP, B) = 2$$
$$u_1\ (P, G, +) = 1$$

and

$$u_2\ (NP, B) = 4$$
$$u_2\ (NP, G, +) = 4$$
$$u_2\ (NP, G, -) = u_2\ (P, G, -) = 3$$
$$u_2\ (P, B) = 2$$
$$u_2\ (P, G, +) = 1\ .$$

The specific numbers used in these definitions of u_1 and u_2 have been picked merely for ease of exposition. The numbers as said, have no significance, and the analysis depends only on the ordinal relationships among them. As a final piece of notation, let the symbol $\oplus$ be used for probability mixtures. More precisely, if u and v are utility numbers and if α satisfies $0 \leqslant \alpha \leqslant 1$, then $\alpha \cdot u \oplus (1 - \alpha) \cdot v$ will stand for the lottery that yields u with probability α and v with probability $1 - \alpha$.

Now it is possible to represent the game being discussed here both in extended and in normal forms:

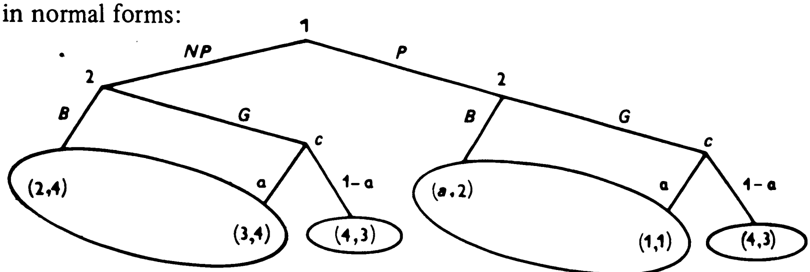

Fig. 1: Extended Form

		Player 2			
		GG	GB	BG	BB
player 1	P	$a\cdot 1+(1-a)\cdot 4$ \ $a\cdot 1+(1-a)\cdot 3$	$a\cdot 1+(1-a)\cdot 4$ \ $a\cdot 1+(1-a)\cdot 3$	a \ 2	a \ 2
	NP	$a\cdot 3+(1-a)\cdot 4$ \ $a\cdot 4+(1-a)\cdot 3$	2 \ 4	$a\cdot 3+(1-a)\cdot 4$ \ $a\cdot 4+(1-a)\cdot 3$	2 \ 4

Fig. 2: Normal Form

If the individual prefers the utility outcome 2 over the utility lottery $\alpha \cdot 1 \oplus (1-\alpha) \cdot 3$, then the strategy *BB* is the dominant strategy for the individual, and the pair (*NP, BB*) is an equilibrium point for $a < 2$ and the pair (*P, BB*) is an equilibrium point for $a > 2$. If the individual prefers the lottery $\alpha \cdot 1 \oplus (1-\alpha) \cdot 3$ to the utility outcome 2, then punishment is a deterrent factor, and the strategy *GB* will be the individual's dominant strategy with the equilibrium point determined by society's preference between the lottery $\alpha \cdot 1 \oplus (1-\alpha) \cdot 4$ and the utility outcome 2. If society is ready to take the risk of incurring injustice the act will be declared as a strict liability offense, and equilibrium will be (*P, GB*). If not, the act will be declared legal, and equilibrium point will be (*NP, GB*).

In all these cases the equilibrum points are also $\max_1 \max_2$ solutions. In other words, they are the outcomes of optimal legal strategies assuming the individual maximizes his utility subject to society's declared strategy. Notice that then the equilibrium points are (*P, GB*) or (*P, BB*) the pair (*NP, GG*) Pareto dominates the equilibrium points.

Society's most preferred pair is (*NP, GG*). However it is not an equilibrium point in the single game. The main message of this paper is that in the repeated game, society can "enforce" this pair of strategies.

3. The Repeated Game

Considering offenses that the individual may repeat many times, we assume that society and the individual both expect that after each play of the game there will be more plays of the same game. The suitable concept in game theory, for analysing this situation, is the repeated game [see for example, *Luce/Raiffa*], which consists of an infinite number of plays of the single game.

This structure seems to be unrealistic. But as *Aumann* [1959] says this notion is more suitable than a repeated game consisting of any fixed number of single games. "The fact that the players know when they have arrived at the last play becomes the decisive factor in the analysis overshadowing all other considerations . . . . *A. W. Tucker* has pointed out the condition that after each play the players expect that there will be more is mathematically equivalent to an infinite sequence of plays".

We assume that each game played at any stage of the repeated game, is identical with every other game, played at any other stage, irrespective of what had happend before.

The individual has complete information after each game. His information set is therefore

$$I_1 = \{\{(NP, B)\}, \{(NP, G, +,)\}, \{(P, G, -)\}, \{(NP, G, -)\}, \{(P, G, +)\}, \{(P, B)\}\}.$$

Since society cannot distinguish between deliberate and accidental acts, its information set at the end of every single game is given by

$$I_2 = \{\{(P, B), (P, G, +)\}, \{(NP, B), (NP, G, +)\}, \{(NP, G, -)\}, \{(P, G, -)\}\}.$$

Both society and the individual have perfect memories. For any natural number k let $I_i(k) = I_i$. Define $I_i^t = \mathop{\times}\limits_{k=1}^{t} I_i(k)$.

In the repeated game, a strategy of a player is a sequence of the form $\{f^t\}_{t=1}^{\infty}$ where $f^1 \in S_1$ and $f^t: I_i^{t-1} \to S_i$. (If $\tau^1, \ldots, \tau^{t-1} \in I_i$ is the information possessed by i, then i will choose the strategy $f^t(\tau^1, \ldots, \tau^{t-1})$. F_i denotes the set of strategies open to i, and $F = F_1 \times F_2$.

Let $h_i^t(f, g)$ be the random variable of the utility of player i at time t, under the assumption that the players adopt strategies $f \in F_1$, $g \in F_2$. For a formal description of $h_i^t(f, g)$ see *Aumann* [1959].

As for the preferences of the players in the repeated game I assume that both society and the individual aim to maximize the limit of the long-run average utility.

If the limit $H_i(f, g) = \lim\limits_{T \to \infty} \frac{1}{T} \sum\limits_{t=1}^{T} h_i^t(f, g)$ exists a.s., the pair (f, g) will be called summable. Denote by $\tilde{F}$ the set of summable pairs of strategies.

I shall write $(f, g) >_i (\bar{f}, \bar{g})$, if there exists $\epsilon > 0$ such that with positive probability there are an infinite number of T for which

$$\frac{1}{T} \sum_{t=1}^{T} h_i^t(f, g) > H_i(\bar{f}, \bar{g}) + \epsilon.$$

Definition: The pair $(f, g) \in \tilde{F}$ will be called an *equilibrium point* of the repeated game if there is no $\bar{f} \in F_1$ or $\bar{g} \in F_2$ satisfying $(\bar{f}, g) >_1 (f, g)$ or $(f, \bar{g}) >_2 (f, g)$.

Let $(f, g) \in \tilde{F}$. The strategy g is said to be $>_2$ maximal relative to f if there is no $\bar{g} \in F_2$ such that $(f, \bar{g}) >_2 (f, g)$.

Now we are ready for the main definiton:

Definition: The pair $(f, g) \in \tilde{F}$ is a *max*$_1$ *max*$_2$ *solution* if g is $>_2$ maximal relative to f and there is no pair $(\bar{f}, \bar{g}) \in \tilde{F}$ such that $\bar{g}$ is $>_2$ maximal relative to $\bar{f}$ and $(\bar{f}, \bar{g}) >_1 (f, g)$.

4. The Theorem

The main theorem asserts the optimality for society of the following legal policy: the individual will be convicted and punished at time $t + 1$ if and only if two conditions hold simultaneously: first the antisocial act is discovered to have been committed at time $t + 1$; and, second, the relative frequency of such acts having occured up to time t is greater than $\alpha + \alpha_t$ where $\{\alpha_t\}$ is a sequence of positive real number converging to 0 "sufficiently slowly". Formally, we have –

Theorem: Let $k > 1$ and $\alpha_t = \sqrt{2k\alpha(1-\alpha)\ln\ln t}\,/\sqrt{t}.$
Let

$\hat{f} = \{\hat{f}^t\}$ be the following social strategy:

$\hat{f}^1 = NP$

$$\hat{f}^{t+1}(\tau^1, \ldots, \tau^t) = \begin{cases} P \text{ if} & \left|\left\{ s \leqslant t \;\middle|\; \begin{array}{l} \tau^s = \{(P, B), (P, G, +)\} \text{ or} \\ \tau^s = \{(NP, B), (NP, G, +)\} \end{array} \right\}\right| \geqslant t(\alpha + \alpha_t) \\ NP \text{ otherwise} \end{cases}$$

Let $\hat{g} = \{\hat{g}^t\}$ be the strategy of the individual where he always adopts G. Then:

a) $H_1(\hat{f}, \hat{g}) = 4 \cdot (1-\alpha) + 3\alpha$

and

$H_2(\hat{f}, \hat{g}) = 3 \cdot (1-\alpha) + 4\,\alpha.$

b) $(\hat{f}, \hat{g})$ is an equilibrium point of the repeated game.

c) $(\hat{f}, \hat{g})$ is a $\max_1 \max_2$ solution.

Proof:

a) Let D_t be the random variable which assumes the value 1 if the chance player causes the act to accur at time t, and assumes the value 0 otherwise.

From the law of the iterated logarithm [see, for example, *Lamperti*] we have that with probability 1 there exists T_0 such that for all $T \geqslant T_0$

$$\frac{1}{T}\sum_{t=1}^{T} D_t - \alpha < \alpha_T .$$

Therefore almost surely society will choose strategy P only a finite numbers of times and with probability 1

$$\frac{1}{T}\sum_{t=1}^{T} h_1^t(\hat{f}, \hat{g}) \to 4 \cdot (1-\alpha) + 3\alpha$$

and

$$\frac{1}{T}\sum_{t=1}^{T} h_2^t(\hat{f}, \hat{g}) \to 3 \cdot (1-\alpha) + 4\,\alpha.$$

b)–c) The ideal combination of strategies from society's point of view is the pair (NP, GG) to be repeated forever. But even this pair will only yield a utility for society of $4 \cdot (1-\alpha) + 3\alpha$. Therefore in order to prove that $(\hat{f}, \hat{g})$ is an equilibrium point in the repeated game and also $\max_1 \max_2$ solution, it suffices to show that there does not exist a $g \in F_2$ with $(\hat{f}, g) >_2 (\hat{f}, \hat{g})$. Let $g \in F_2$ and let $\epsilon > 0$. We now show the emptiness of the event "for infinitely many T,

$$\frac{1}{T}\sum_{t=1}^{T} h_2^t(\hat{f}, g) > H_2(\hat{f}, \hat{g}) + \epsilon".$$

Let $\{w_t\}$ be a sequence of the individual's utilities obtained from $(\hat{f}, g)$. Denote $\bar{w}_T = \frac{1}{T}\sum_{t=1}^{T} w_t$. Let N be the minimal natural number for which $\alpha_N + 4/N < \epsilon$. Our claim is that only for a finite number of times T, $\bar{w}_T > 4\alpha + 3(1-\alpha) + \epsilon$. The claim is applied easily from the following three assertions:

For any $t > N$:

1. if $\bar{w}_t > 4(\alpha + \alpha_N) + 3(1 - \alpha - \alpha_N)$ then $\bar{w}_{t+1} < \bar{w}_t - \alpha/(t+1)$.
2. if $\bar{w}_t \leqslant 4(\alpha + \alpha_N) + 3(1 - \alpha - \alpha_N)$ then $\bar{w}_{t+1} < 4\alpha + 3(1-\alpha) + \epsilon$.
3. there is $T > N$ such that $\bar{w}_T \leqslant 4(\alpha + \alpha_N) + 3(1 - \alpha - \alpha_N)$.

Proof of assertions 1:

For any $t > N$ if $\bar{w}_t > 4(\alpha + \alpha_N) + 3(1 - \alpha - \alpha_N)$ then also

$$\bar{w}_t > 4(\alpha + \alpha_t) + 3(1 - \alpha - \alpha_t).$$

The relative frequency of the forbidden acts is greater than $\alpha + \alpha_t$ and therefore society's strategy at time $t+1$ is P.

Now $w_{t+1} \leqslant 3$ and

$$\bar{w}_{t+1} \leqslant \frac{t \cdot \bar{w}_t + 3}{t+1} \leqslant \bar{w}_t + \frac{3 - \bar{w}_t}{t+1} \leqslant \bar{w}_t - \frac{\alpha + \alpha_t}{t+1} < \bar{w}_t - \frac{\alpha}{t+1}$$

Proof of assertion 2:

If $\bar{w}_t \leqslant 4(\alpha + \alpha_N) + 3(1 - \alpha - \alpha_N)$ then

$$\bar{w}_{t+1} \leqslant \frac{[4(\alpha + \alpha_N) + 3(1 - \alpha - \alpha_N)]\,t + 4}{t+1} = [4\alpha + 3(1-\alpha)]\frac{t}{t+1} + \frac{t \cdot \alpha_N + 4}{t+1} < 4\alpha + 3(1-\alpha) + \alpha_N + \frac{4}{N} < 4\alpha + 3(1-\alpha) + \epsilon.$$

Proof of assertion 3: The harmonie series diverges and therefore assertion 1) implies 3.

Remark: If we replace the sequence $\{\alpha_t\}$ with the sequence $\alpha_t \equiv \epsilon > 0$, the previously optimal strategy is no longer optimal since the individual may deviate with relative frequency of say $\epsilon/2$, which would result almost surely in being punished only a finite number of times.

If society convicts the accused whenever the frequency of the harmful act exceeds α (i.e. if we let $\alpha_t \equiv 0$) then the individual will be punished infinitely many times (almost surely) even if he consistently adopts the stragey G. The same result is true if we choose $k < 1$ in the definition of $\{\alpha_t\}$ [see *Lamperti*]. However any sequence that tends to 0 more slowly than our sequence is suitable for a definition of an optimal strategy.

Remark: Some of the assumptions made above may be dropped or relaxed. For example, we can remove our assumption that committal of a forbidden act is always found out by society if we introduce a chance player representing the chance of discovery. Society's information set is represented here by loops.

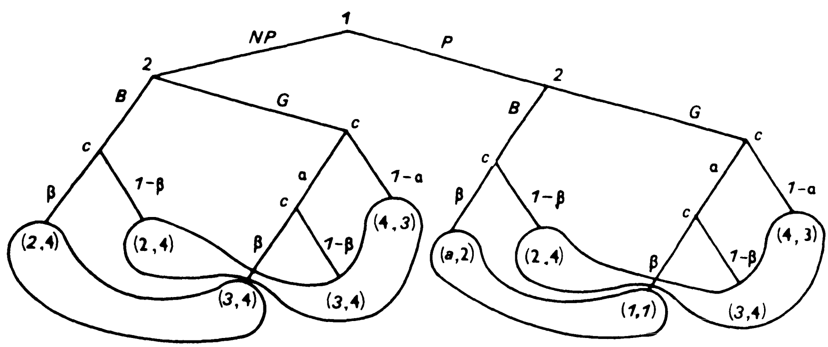

Fig. 3

In a way similar to the main theorem we can prove that there is a positive sequence $\beta_n \to 0$ and an optimal policy which punishes the individual if the frequency of his offences in the past is greater than $\alpha \cdot \beta + \beta_n$, where β is the probability of discovery.

References

Aumann, R.J.: Acceptable Points in General Cooperative n-person Game. Ed. by A.W. Tucker and R.C. Luce, Contributions to the Theory of Games IV. Annals of Math. Studies, **40**, Princeton, N.J., 1959, 287–324.

Lamperti, J.: Probability. New York 1966.

Luce, R.P., and *H. Raiffa*: Games and Decisions. New York 1957.