

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Essays in the Economics of Crime and Punishment

Volume Author/Editor: Gary S. Becker and William M. Landes, eds.

Volume Publisher: NBER

Volume ISBN: 0-87014-263-1

Volume URL: <http://www.nber.org/books/beck74-1>

Publication Date: 1974

Chapter Title: The Optimum Enforcement of Laws

Chapter Author: George J. Stigler

Chapter URL: <http://www.nber.org/chapters/c3626>

Chapter pages in book: (p. 55 - 67)

# The Optimum Enforcement of Laws

George J. Stigler

*University of Chicago and National Bureau of Economic Research*

All prescriptions of behavior for individuals require enforcement. Usually the obligation to behave in a prescribed way is entered into voluntarily by explicit or implicit contract. For example, I promise to teach certain classes with designated frequency and to discuss matters which I, and possibly others, believe are relevant to the course titles. By negotiation, and in the event of its failure, by legal action, I and my employer seek to enforce the contract of employment against large departures from the promised behavior. Performance of some kinds of behavior is difficult or impossible to enforce—such as promises to be creative, noble, or steadfast in crisis—and as a result such contractual promises are either not made or enforced only when there is an uncontroversially flagrant violation. The influence upon contract, and upon economic organization generally, of the costs of enforcing various kinds of contracts has received virtually no study by economists, despite its immense potential explanatory power.

When the prescribed behavior is fixed unilaterally rather than by individual agreement, we have the regulation or law, and enforcement of these unilateral rules is the subject of the present essay. Departures of actual from prescribed behavior are crimes or violations, although one could wish for a less formidable description than “criminal” to describe

many of the trifling offenses or the offenses against unjust laws. My primary purpose is to construct a theory of rational enforcement, a theory which owes much to Gary Becker's major article on the subject (1968). In the conclusion the problem of explanation, as distinguished from prescription, will be commented upon.

## I. THE GOAL OF ENFORCEMENT

The goal of enforcement, let us assume, is to achieve that degree of compliance with the rule of prescribed (or proscribed) behavior that the society believes it can afford. There is one decisive reason why the society must forego "complete" enforcement of the rule: enforcement is costly.

The extent of enforcement of laws depends upon the amount of resources devoted to the task. With enough policemen, almost every speeding automobile could be identified. The success of tenacious pursuit of the guilty in celebrated crimes (such as the great English train robbery and the assassination of Martin King) suggests that few crimes of sane men could escape detection. We could make certain that crime does not pay by paying enough to apprehend most criminals. Such a level of enforcement would of course be enormously expensive, and only in crimes of enormous importance will such expenditures be approached. The society will normally give to the enforcement agencies a budget which dictates a much lower level of enforcement.

The cost limitation upon the enforcement of laws would prevent the society from forestalling, detecting, and punishing all offenders, but it would appear that punishments which would be meted out to the guilty could often be increased without using additional resources. The offender is deterred by the expected punishment, which is (as a first approximation) the probability of punishment times the punishment—\$100 if the probability of conviction is 1/10 and the fine \$1,000. Hence, increasing the punishment would seem always to increase the deterrence. Capital punishment is cheaper than long term imprisonment; and seizure of all the offender's property may not be much more expensive than collecting a more moderate fine.

To escape from this conclusion, Becker introduces as a different limitation on punishment the "social value of the gain to offenders" from the offense. The determination of this social value is not explained, and one is entitled to doubt its usefulness as an explanatory concept: what evidence is there that society sets a positive value upon the utility derived from a murder, rape, or arson? In fact the society has branded the utility

derived from such activities as illicit. It may be that in a few offenses some gain to the offender is viewed as a gain to society,<sup>1</sup> but such social gains seem too infrequent, small, and capricious to put an effective limitation upon the size of punishments.

Instead we take account of another source of limitation of punishment, which arises out of the nature of the supply of offenses. It is no doubt true that the larger the punishment, the smaller will be the expected net utility to the prospective offender from the commission of a given offense. But marginal decisions are made here as in the remainder of life, and the marginal deterrence of heavy punishments could be very small or even negative. If the offender will be executed for a minor assault and for a murder, there is no marginal deterrence to murder. If the thief has his hand cut off for taking five dollars, he had just as well take \$5,000. Marginal costs are necessary to marginal deterrence.<sup>2</sup> The marginal deterrence to committing small crimes is also distorted if an otherwise appropriate schedule of penalties is doubled or halved.

One special aspect of this cost limitation upon enforcement is the need to avoid overenforcement. The enforcement agency could easily apprehend most guilty people if we placed no limits upon the charging and frequent conviction of innocent people. In any real enforcement system, there will in fact be conviction and punishment of some innocent parties, and these miscarriages of justice impose costs of both resources and loss of confidence in the enforcement machinery. The costs of defense of innocent parties, whether borne by themselves or by the state, are part of the costs of enforcement from the social viewpoint. The conviction of innocent persons encourages the crime because it reduces the marginal deterrence to its commission.

The significance of an offense to society—the quantity of resources that will be used to “prevent” the offense—will in general increase with the gravity of the offense. The increase in resources, however, will not manifest itself only in an increase in punishments. The state will pursue more tenaciously the offender who commits a larger crime (or repeti-

---

1. For example, the thief reduces the welfare expenditures of the state, or the arsonist warms the neighboring houses.

2. Becker writes the expected utility from an offense as  $EU = pU(Y-f) + (1-p)U(Y)$ , where  $Y$  is the money value of the gain,  $p$  the probability of detection and conviction, and  $f$  the fine. (The income,  $Y$ , and fine,  $f$ , must be interpreted as average annual flows; for a single offense  $Y$  must be less than  $f$ .) If this expression is differentiated partially with respect to  $Y$ ,  $\delta EU/\delta Y = pU'(Y-f) + (1-p)U'(Y) = U'(Y) - pfU''(Y)$ , which is positive for all  $Y$ . Of course  $p$  and  $f$  will increase with  $Y$  to prevent this incitement to larger crime.

tive crimes) and thus increase also the probability of apprehending him.

There is a division of labor between the state and the citizen in the prevention of virtually every offense. The owner of large properties is required to do much of his direct policing: there are surely more watchmen and guards than policemen in a typical city. The larger accumulations of wealth, moreover, are to be guarded by the owner through devices such as nonnegotiability and custody of funds by specialists. Accordingly, the *public* punishments for crimes against property do not increase in proportion to the value of the property. In the protection of people, as distinguished from their property, the individual is required to protect himself from minor offenses or at least to detect their occurrence and assume a large part of the burden of prosecution (for example, shoplifting, insults, simple trespass), but he is allowed less discretion in prosecution for major assaults.

The relationship of duration and nature of penalties to age and sex of offender, frequency of previous offenses, and so forth is also explicable in terms of cost of enforcement. The first-time offender may have committed the offense almost accidentally and (given any punishment) with negligible probability of repetition, so heavy penalties (which have substantial costs to the state) are unnecessary. The probability of a repetition of an offense by a seasoned offender is also zero during his imprisonment, so the probability of repetition of an offense is relevant to the penalty also in his case.

Indeed, the problem of determining the efficient penalty may be viewed as one in statistical inference: to estimate the individual's average, durable propensity to offend (the population value) on the basis of a sample of his observed behavior and how this propensity responds to changes in penalties. As in other sequential sampling problems, one can estimate this propensity more accurately, the longer the individual's behavior is observed.

The society will be more concerned (because each individual is more concerned) with major than minor offenses in the following sense: there is increasing marginal disutility of offenses, so a theft of \$1,000 is more than twice as harmful as a theft of \$500. In the area of offenses to property, this result is implied by diminishing marginal utility of income. In the area of offenses to persons, it is more difficult to measure damage in any direct way, but a similar rule probably holds.

So much for prevention and punishment; let us turn to the offenses.

## II. THE SUPPLY OF OFFENSES

The commission of offenses will be an act of production for income or an act of consumption. A consumption offense would be illustrated by speeding in an automobile used for recreation or assaulting a courtship rival (when the girl is poor). A production offense would be illustrated by theft, smuggling, and the violation of economic regulations. In the realm of offenses to property, income objectives are of course paramount, and we may recall Adam Smith's emphasis upon the economic nature of crime:

The affluence of the rich excites the indignation of the poor, who are often both driven by want, and prompted by envy, to invade his possessions. It is only under the shelter of the civil magistrate that the owner of that valuable property, which is acquired by the labour of many years, or perhaps of many successive generations, can sleep a single night in security. He is at all times surrounded by unknown enemies, whom, though he never provoked, he can never appease, and from whose injustice he can be protected only by the powerful arm of the civil magistrate continually held up to chastise it. The acquisition of valuable and extensive property, therefore, necessarily requires the establishment of civil government. Where there is no property, or at least none that exceeds the value of two or three days labour, civil government is not so necessary (Smith, 1937, p. 670).

The professional criminal seeks income, and for him the usual rules of occupational choice will hold. He will reckon the present value of the expected returns and costs of the criminal activity and compare their difference with the net returns from other criminal activities and from legitimate activities. The costs of failure in the execution of the crime correspond to the costs of failure in other occupations. The costs of injuries to a professional athlete are comparable to the costs to the offender of apprehension, defense, and conviction, but normally legal occupations have only monetary costs of failure.

The details of occupational choice in illegal activity are not different from those encountered in the legitimate occupations. One must choose the locality of maximum income expectation (and perhaps, like a salesman, move from area to area). One must choose between large, relatively infrequent crimes and smaller, more frequent crimes. One must reckon in periods of (involuntary) unemployment due to imprisonment. Earnings can be expected to rise for a time with experience.

The probability of apprehension (and therefore of conviction) is an increasing function of the frequency of commission of offenses. If the probability of detection is  $p$  for one offense, it is  $1 - (1 - p)^n$  for at least one conviction in  $n$  offenses, and this expression approaches unity as  $n$  becomes large. In fact, the probability of detection ( $p$ ) rises after each apprehension because the enforcement agency is also learning the offender's habits. On this score alone, there is a strong incentive to the criminal to make very infrequent attempts to obtain very large sums of money. The probability of success is also affected by the precautions of the prospective victim: Fort Knox is more difficult to enter than a liquor store. The efforts of detection will also increase with the size of the offense.

We may postulate, in summary, a supply of offenses which in equilibrium has the following properties:

1. Net returns are equalized, allowance being made for risk and costs of special equipment required for various activities.
2. The determinants of supply which are subject to the control of society are: (a) the structure of penalties by offense; (b) the probability of detection for each offense; (c) certain costs of the conduct of the offending activity; for example, the cost of making successful counterfeit money can be increased by complicating the genuine money.
3. The penalties and chances of detection and punishment must be increasing functions of the enormity of the offense.

Although it smacks of paradox, it may be useful to reinterpret the offending activity as providing a variety of products (offenses). These offenses are in a sense demanded by the society: my wallet is an invitation to the footpad, my office funds to the embezzler. The costs of production of the offenses are the ordinary outlays of offenders plus the penalties imposed by the society. The industry will operate at a scale and composition of output set by the competition of offenders and the cost of producing offenses.

The structure of rational enforcement activities will have these properties:

1. Expected penalties increase with expected gains so there is no marginal net gain from larger offenses. Let the criminal commit in a year  $S$  crimes of size  $Q$ , where  $Q$  is the monetary value to the criminal of the successful completion of the crime. The fraction

( $p$ ) of crimes completed successfully (or the probability of successful completion of one crime) is a decreasing function of the amount of expenditure ( $E$ ) undertaken by society to prevent and punish the crime. (Punishment is used for deterrence, and is only a special form of prevention.) Hence  $p = p(E, Q)$ , or possibly  $p = p(E, Q, S)$ . The expected punishment is the fraction of crimes apprehended (and punished) times the punishment,  $F$ . The condition for marginal deterrence is, for all  $Q$ ,  $d(pSQ)/dQ \leq d(1 - p)SF/dQ$ .

2. The expenditures on prevention and enforcement should yield a diminution in offenses, at the margin, equal to the return upon these resources in other areas. An increment of expenditures yields a return in reduced offenses,

$$\sum_{Q'} d(pSQ')/dE = \text{marginal return on expenditures elsewhere,}$$

where  $Q'$  is the monetary value of the offense to society.<sup>3</sup>

I do not include foregone lawful services of the criminal in the cost of his activity to society (= noncriminals) since he, not others, would receive the return (taxes aside!) if he shifted from crime to a lawful occupation.

### III. THE ENFORCEMENT AGENCY: A NORMATIVE APPROACH

A law is enforced, not by "society," but by an agency instructed to that task. That agency must be given more than a mandate (an elegant admonition) to enforce the statute with vigor and wisdom: it must have incentives to enforce the law efficiently. There are at least two deficiencies in the methods by which most agencies are induced to enforce the laws properly.

The first deficiency is that the enforcement agency does not take into account, at least explicitly and fully, the costs it imposes upon the activity or persons regulated. In the area of ordinary criminal offenses, the society will, if anything, wish to increase (at no expense to itself) the costs of defense for guilty persons, but it should not impose costs (and certainly not unnecessary costs) upon innocent parties. In fact, the administration of criminal justice should in principle include as a cost the reimbursement

---

3. Currency has the same value to the criminal as to society, so  $Q' = Q$ . But for any commodity which does not have a market price independent of ownership,  $Q < Q'$ .



of the expenses of defense of people charged and acquitted. The compensation actually paid will not exactly compensate injured persons, because of the administrative costs of ascertaining exact compensation, but the taking of an innocent person's personal wealth, including foregone income, differs in no respect from the taking of some of his real estate (for which under eminent domain it is necessary to compensate him fully).<sup>4</sup>

In the area of economic regulation, guilt is often an inappropriate notion, and when it is inappropriate all costs of compliance must be reckoned into the social costs of enforcement. The utility's costs in preparing a rate case or Texas Gulf Sulfur's costs in defending itself against the Securities and Exchange Commission are social costs of the regulatory process. Reimbursement is now achieved by charging the consumers of the products and the owners of specialized resources of these industries: they bear the private costs of the regulatory process. This is at least an accidental allocation of costs, and when the regulation seeks to aid the poorer consumers or resource owners, a perverse allocation.

The second deficiency in the design of enforcement is the use of inappropriate methods of determining the extent of enforcement. The annual report of an enforcement agency is in effect the justification of its previous expenditures and the plea for enlarged appropriations. The Federal Trade Commission will tell us, for example, that in fiscal 1966 as part of its duty to get truthful labeling of furs and textiles, it inspected 12,625 plants and settled 213 "cases" for \$1,272,000 plus overhead. The agency may recite scandals corrected or others still unrepressed, but it neither offers nor possesses a criterion by which to determine the correct scale of its activities.

A rational measure of enforcement procedure could in principle be established in almost any area. Consider the fraudulent labeling of textiles. We could proceed as follows:

1. The damage to the consumer from the purchase of a mislabeled textile could be estimated, and will obviously vary with the mislabeling (assuming that the legal standards are sensible!). The difference between market value of the true and alleged grades is one component of the damage. A second and more elusive component is the additional cost of deception (earlier replacement, skin irritation, and so forth): the consumer who would not have purchased the inferior quality at a competitive price had he known its

---

4. It is an interesting aspect of our attitudes in this area that many people believe that acquitted persons are probably guilty.

inferiority has suffered additional damage. Thus the measure of damage is the amount a consumer would pay to avoid the deception, that is, the value of the *insured* correct quality minus the value of the actual quality.

2. As a matter of deterrence, the penalties on the individual mislabeler should be equal to a properly taken sum of the following items: (a) The damage per yard times the number of yards, say per year. Let this be  $H$ . (b) The costs of the enforcement agency, say  $E$ , per year. This sum should include reimbursement of the costs of those charged and acquitted. (c) The costs of defense (if detected) for the mislabeler,  $D$ . Where guilt is an appropriate notion, as presumably in the case of mislabeling, the society may wish to ignore these costs, which is to say, resources devoted to this end deserve no return. Where guilt is inappropriate, these costs should be reckoned in. The sum of these penalties must be multiplied by  $1/p$ , where  $p$  is the probability of detection of the offense within the year. This probability is a function of  $E$  and  $H$ .
3. The enforcement agency should minimize the sum of damages plus enforcement costs,  $(\Sigma h + E)$  or  $(\Sigma H + \Sigma D + E)$ .<sup>5</sup>

This goal will serve two functions. The first is to set the scale of enforcement, namely where marginal return equals marginal cost. If the scale of enforcement is correct, society is not spending two dollars to save itself one dollar of damage, or failing to spend one dollar where it will save more than that amount of damage. The second function is to guide the selection of cases: the agency will not (as often now) seek numerous, easy cases to dress up its record, but will pursue the frequent violator and the violator who does much damage.

This sort of criterion of enforcement is readily available in certain areas. The secret service, for example, reports that in fiscal 1967 the loss to the public from counterfeit money was \$1,658,100.75 (an excellent instance of counterfeit accuracy). Perhaps half of the \$17 million spent by this agency was devoted to the suppression of counterfeiting, and to this one must add the costs of legal actions, imprisonment, and so forth. The secret service should be asking whether the amount of counterfeit money passed would fall by a dollar if a dollar more were spent on enforcement costs minus the corresponding fines collected.

---

5. The fines will be  $\Sigma H/p$ , but the fines per se are transfers rather than social costs; see Becker (1958), pp. 180-81.

The penalty structure should incorporate the social appraisal of the importance of the suppression of the offenses. The law does not in general provide this scale of values, as can be shown by the list of maximum penalties for the violations of economic regulations listed in Table 1.

The use of criminal sanctions is erratic, and the implicit equivalence of fines and imprisonment varies from \$1,000 per year to \$10,000 per

TABLE 1  
PUBLIC PENALTIES FOR VIOLATION OF ECONOMIC STATUTES

Offense	Enforcement Agency	Maximum Penalty	Statute
Restraint of trade	Antitrust Division	\$50,000 + 1 year (+ triple damages and costs)	Sherman (1890, 1955)
Unfair methods of competition	FTC	Cease and desist order	FTC (1914)
Refusal to testify, or testify falsely, under same	FTC	\$1,000-\$5,000 + 1 year	Same
Price discrimination	FTC	Cease and desist (+ triple damages and costs)	Clayton (1914)
False advertisements of foods, drugs, or cosmetics	FTC	\$5,000 + 6 months	Wheeler-Lea (1938)
Adulteration or misbranding of food	Secretary of Health, Education, and Welfare	\$1,000 + 1 year, first offense; \$10,000 + 3 years, later offense	Copeland Act (1938)
Exporting apples and pears without certificate of quality	Dept. of Agriculture	Denial of certificate for 10 days; \$100-\$1,000 for knowing violation	Apples and Pears for Export (1933)
Exporting apples in improper barrels	Dept. of Agriculture	\$1 barrel	Standard Barrels and Standard Grades of Apples Act (1912)

TABLE 1 (Concluded)  
PUBLIC PENALTIES FOR VIOLATION OF ECONOMIC STATUTES

Offense	Enforcement Agency	Maximum Penalty	Statute
Exporting other fruit or vegetables in improper barrels	Dept. of Agriculture	\$500 or 6 months if willful	Standard Barrels . . . for Fruits, Vegetables and Other Dry Commodities (1915)
Exporting grapes in improper baskets	Dept. of Agriculture	\$25	Standard Baskets Act (1916)
Giving rebates in freight charges (trucks)	ICC	\$200-\$500 first offense; \$250-\$5,000 repeated offense	Motor Carrier Act (1935)
Same, water carriers	ICC	\$5,000 if willful	Transportation Act (1940)
Failure to disclose interest charges	FRB	Twice finance charge, within \$100-\$1,000; \$5,000 and/or 1 year if willful	Consumer Credit Protection Act (1968)
Falsely certify a check	FBI	\$5,000 and/or 5 years	62 Stat. 749 (June 25, 1948)
Failure to deliver gold or certificates to FR Bank when ordered	FRB	Twice the number of dollars	Federal Reserve Act (1913)
Evasion of excise taxes	Treasury	\$10,000 and/or 5 years if willful; forfeiture of goods and conveyance	Revenue Act (1954)
Securities Act violation	SEC	\$5,000 and/or 5 years	Securities Act (1934)
Misbrand hazardous substances	Secretary of Health, Education, and Welfare	\$500 and 90 days; \$3,000 or 1 year if willful or repeated	Hazardous Substances Act (1960)

year. Many of the penalties are not even stated in the statutes: the penalty for drinking (industrial) alcohol which has not paid its beverage tax is sometimes blindness or death. The penalties for the essentially similar offense (if such it must be called) of reducing freight rates can be ten times as much for a barge operator as for a trucker. Of course these maximum penalties are not actual penalties, but one is not entitled to hope for much more rationality or uniformity in the fixing of penalties for specific offenses. (The lawyers have apparently not studied in adequate detail the actual sanctions for economic offenses.)

One may conjecture that two features of punishment of traditional criminal law have been carried over to economic regulation: the attribution of a substantial cost to the act of conviction itself, and the related belief that moral guilt does not vary closely with the size of the offense. Whatever the source, the penalty structure is not well designed for either deterrence or guidance of enforcement.

#### IV. CONCLUSION

The widespread failure to adopt rational criteria of enforcement of laws has been due often and perhaps usually to a simple lack of understanding of the need for and nature of rational enforcement. The clarification of the logic of rational enforcement, and the demonstration that large gains would be obtained by shifting to a rational enforcement scheme, are presumably the necessary (and hopefully sufficient) conditions for improving public understanding of enforcement problems.<sup>6</sup>

There is, however, a second and wholly different reason for the use of what appear to be inappropriate sanctions and inappropriate appropriations to enforcement bodies: the desire of the public *not* to enforce the laws. The appropriations to the enforcement agency and the verdicts of juries are the instruments by which the community may constantly review public policy. If the society decides that drinking alcoholic beverages or speeding in automobiles is not a serious offense in its ordinary form, they may curtail resources for enforcement and so compel the enforcement agency to deal only with a smaller number of offenses (perhaps offenses of larger magnitude, such as chronic drunkenness or driving at extremely high speeds). There is considerable inertia in the legislative

---

6. The peculiarities of the structure of sanctions in economic regulation are partially due also to the response of the regulated businesses. They may effectively lobby to limit appropriations to the regulatory body, but they can also impose costly activities upon the regulatory body which force it to curtail other controls.

process — inertia that serves highly useful functions — and it is much easier to make continuous marginal adjustments in a policy through the appropriations committee by varying the resources for its enforcement than it is to modify the statute. Variation in enforcement provides desirable flexibility in public policy.

#### REFERENCES

- Becker, Gary. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (March–April 1968). Included in this volume.
- Smith, Adam. *The Wealth of Nations*. New York: Modern Library, 1937.