ELSEVIER International Review of Law and Economics 24 (2004) 209–218

International Review of Law and Economics

# The advantage of focusing law enforcement effort

Henrik Lando[a], Steven Shavell[b,*]

[a] *Copenhagen Business School, Denmark*
[b] *Harvard Law School, Cambridge, MA 02138, USA*

**Abstract**

It is shown in this paper that there may exist an intrinsic advantage in focusing law enforcement effort on a subgroup of possible violators of law, rather than applying law enforcement effort uniformly over the relevant population of potential violators. For example, it may be desirable for the tax audit rate to be higher in one region of the country than another. This may be desirable even though, as is assumed, the frequency of violations does not differ among regions.


*JEL classification:* D8, H00; K00; K14; K42

*Keywords:* Optimal enforcement; Focused enforcement; Allocation of police

## 1. Introduction

In this paper, we demonstrate and develop the point that there may exist an intrinsic advantage in focusing law enforcement effort on a subgroup of possible violators of law, rather than applying law enforcement effort uniformly over the relevant population of potential violators. For example, it may be desirable for tax examiners to undertake more audits of individuals who live in one region of the country than another; or it may be desirable for police to patrol one highway more intensively than another. This may be desirable even though, as we assume, the frequency of violations is the same throughout the population of potential violators—even though the likelihood of tax violations is the same in one region as in another, or even though the propensity to speed

* Corresponding author. Tel.: +1 617 495 3668; fax: +1 617 496 2256.
*E-mail address:* shavell@law.harvard.edu (S. Shavell).

0144-8188/$ – see front matter © 2004 Published by Elsevier Inc.
doi:10.1016/j.irle.2004.08.005

is the same among those who drive on one highway as it is among those who drive on another.[1]

The kernel of the argument to be made is that it may well be the case that the gain in deterrence in the group for which enforcement effort is focused outweighs the loss in deterrence in the group for which enforcement effort is lowered. To illustrate, suppose that each individual who files a tax return will obtain a benefit of \$100 if he falsely claims a deduction and that the penalty for making a false claim is \$1000. Suppose too that, if the available staff of tax examiners audits all individuals uniformly, the audit rate will be 6%. Then, since the expected penalty will be only \$60, all individuals will make false claims in order to gain \$100 (assume that they are risk neutral)—the auditing will be ineffective. However, if the tax authorities announce that they will devote all of their staff to audit just the taxpayers living in one half of the country with probability 12%, these individuals will be deterred from making false claims by the \$120 expected penalty, so deterrence will rise from none to perfect deterrence for them. For the other taxpayers, the audit probability will fall from 6% to zero, but deterrence will be unaffected—it will remain at the level zero. Hence, the focusing of law enforcement effort on half of the taxpayers raises deterrence overall and is desirable. In this illustration our claim is true because, among other things, all individuals obtain the same benefit from a violation, and a uniform law enforcement policy deters none. But our point holds under broad circumstances, as we are now about to describe.

In Section 2 of the paper, we analyze formally the general question of how best to allocate a given level of law enforcement resources across a population of potential violators.[2] The question that we address is thus different from, and complementary to, that initially posed by Becker (1968). Becker asked about the optimal level of resources to employ in enforcing law, but not about the allocation of these resources. We demonstrate that if the available enforcement resources are below a critical threshold, the optimal enforcement policy is to focus the *entire* enforcement resources in a subgroup of the population, and thus not to enforce the law at all elsewhere. When the available enforcement resources are above the critical threshold, it will still sometimes be desirable to enforce the law with different intensities for one subgroup from that for another, but the lesser intensity will be positive (rather than zero). Also, we show that it will never be desirable to divide the population into more than two subgroups for purposes of law enforcement. This latter result simplifies the determination of the optimal law enforcement policy.

In Section 3, we consider briefly several factors that bear on our analysis and its interpretation.

## 2. Model and analysis

Risk-neutral individuals living in a city might decide to violate a legal rule, depending on the benefits that they would obtain from doing so and the probability and magnitude of sanc-

[1] We are not making the obvious point that if the frequency of violations is higher for some subgroup (such as tax filers who are self-employed), that it may then be desirable to devote more law enforcement resources to that subgroup.

[2] As far as we know, this question has not been studied before.

tions they would face for a violation. In particular we assume that, at each location $t \in [0, 1]$ in the city, the probability distribution of benefits that individuals who reside there would obtain from a violation is the same; that is, there is no difference in the subpopulation of individuals living at one location from that living at another location.[3] Let this distribution be as follows:

$b$ = benefit that a person obtains if he violates the legal rule; $b \in [0, m]$.
$f(b)$ = probability density of $b$.

For simplicity, we assume that the population density is the same at all locations/and that the total population size is 1. A violation of the rule causes harm:

$h$ = harm from a violation; $h > 0$.
If a violator is caught, he will suffer a sanction:
$s$ = sanction; $s \geq 0$.

The sanction is monetary and socially costless to impose.[4] We assume for simplicity that $s$ is fixed, for our interest is in the optimal allocation of enforcement resources given any $s$.

The probability of sanctions at location $t$ is assumed to be proportional to the police enforcement resources employed at that location (the number of people that can be checked by two policemen is twice the number that can be checked by one policeman). There is a limited level of enforcement resources available in the city, which is represented by the probability with which individuals in the city would be caught if enforcement effort were uniform at different locations across the city. Let

$P$ = probability of sanctions for violations if law enforcement is uniform across locations.

An *enforcement policy* corresponds to a function giving the probability of enforcement at each location $t$ in the city:[5]

$p(t)$ = probability of enforcement at location $t$; $0 \leq p(t) \leq 1$, where $t \in [0, 1]$.

A policy $p(t)$ is presumed to be piecewise continuous with no isolated points.[6] A policy must obey the resource constraint that:[7]

$$\int_0^1 p(t)\,\mathrm{d}t = P. \tag{1}$$

Let us now describe individuals' behavior and social welfare given an enforcement policy. At a location $t$ in the city, the expected sanction is $p(t)s$, so that a person living at $t$ will violate the legal rule if and only if:[8]

$$b > p(t)s. \tag{2}$$

[3] The assumption that the individuals live in a city is used only for concreteness. More generally, the individuals are indexed by some parameter (such as the highway on which they drive) which plays the role of location in the city. Thus, our assumption is that the density of benefits is the same regardless of the value of the parameter.

[4] Our main results would apply as well if the sanction is socially costly to impose.

[5] It will be evident that our results do not depend on the assumption that $t$ is continuously variable rather than discrete, or that it is a scalar rather than a vector.

[6] More precisely, $p(t)s$ is continuous either from the left or from the right.

[7] We assume the resource constraint is binding for simplicity. In fact, the constraint will be binding *if* $Ps < h$, which can be shown to be true if $P$ is optimally determined; see, for example, Polinsky and Shavell, 2000.

[8] We thus assume that if $b = p(t)s$, the individual is deterred; this assumption is made for convenience.

Hence, social welfare at location $t$ is:

$$w(p(t)) = \int_{p(t)s}^{m} (b-h)f(b)\,\mathrm{d}b \tag{3}$$

and social welfare in the city is:

$$\int_{0}^{1} w(p(t))\,\mathrm{d}t. \tag{4}$$

The social problem is to choose an enforcement policy $p(t)$ to maximize social welfare (4) subject to the constraint (1) on enforcement resources. The optimal enforcement policy will be denoted by $p^*(t)$.[9] To solve this problem, we first state an important fact about the solution.

**Proposition 1.** *There exists an optimal enforcement policy involving at most two different probabilities.That is,there exists a* $p^*(t)$*such that* $p^*(t)= p_1$ *in some* $X \subset [0, 1]$ *and* $p^*(t)= p_2$ *in* $Y = [0, 1] \sim X$.

Note that $p_1$ may equal $p_2$, so that Proposition 1 allows for the possibility of a uniform optimal probability. The proposition is proved in the Appendix. The proposition is first shown to hold for policies $p(t)$ that takes on only a finite number of probabilities. The argument is, in essence, that any such policy under which there exist three regions with three different probabilities can be replaced with a policy under which the three regions are converted into two regions with two different probabilities without reducing social welfare. Then a limiting argument demonstrates that, since the proposition holds for finite-valued policies, it must hold for piecewise-continuous policies.[10]

Given Proposition 1, the determination of the optimal policy $p^*(t)$ devolves into a straightforward problem. The two regions $X$ and $Y$ can be taken to be intervals, as social welfare clearly is not affected by interchanging enforcement effort at one location with that at another (or equivalently, by a relabeling of locations). Hence, we need only consider pairs of probabilities $p_1$ and $p_2$ over two regions of length $k$ and $1-k$, where $k \in [0, 1]$, obeying the enforcement resource constraint (1), namely,

$$kp_1 + (1-k)p_2 = P. \tag{5}$$

That is, the social problem is to maximize social welfare (4), which reduces to:

$$kw(p_1) + (1-k)w(p_2), \tag{6}$$

[9] The optimal policy $p^*(t)$ is determined only up to a set of measure zero; obviously, a policy that differs from $p^*(t)$ at several points would also be optimal. We will not bother to qualify our statements about the optimal $p^*(t)$ in this regard below.

[10] The proposition does not follow from the Lagrangean condition for maximizing (4) subject to (1), namely, from the condition that $w'(p(t)) = \lambda$.

subject to (5). Since (5) implies that $p_2 = (P - kp_1)/(1 - k)$, the social problem is simply to maximize:

$$kw(p_1) + (1 - k)((P - kp_1)/(1 - k)) \tag{7}$$

over $k$ and $p_1$.

Before continuing, let us define some additional notation. Let

$$v(p) = w(p) - w(0) = \int_{ps}^{m} (b - h)f(b)\,\mathrm{d}b - \int_{0}^{m} (b - h)f(b)\,\mathrm{d}b = \int_{0}^{ps} (h - b)f(b)\,\mathrm{d}b. \tag{8}$$

In other words, $v(p)$ is the increase in social welfare at any location due to enforcing with probability $p$ rather than not enforcing at all; the increase in social welfare is associated with a net social benefit of $(h - b)$ for all individuals who are deterred by $ps$ who would have obtained gains from violating less than or equal to $ps$. Let

$$y(p) = \frac{v(p)}{p} = \frac{\int_{0}^{ps} (h - b)f(b)\,\mathrm{d}b}{p} \tag{9}$$

be the gain in social welfare per unit of enforcement effort, that is, the gain in social welfare per policeman, when $p$ is the probability of sanctions. It is informative to examine the derivative of (9):

$$y'(p) = \frac{sp(h - ps) - \int_{0}^{ps} (h - b)f(b)\,\mathrm{d}b}{p^2}. \tag{10}$$

Thus, raising the probability of enforcement increases the social return per policeman only if $sp(h - ps)f(ps)$ exceeds the integral, which is to say (dividing by $p$), only if $s(h - ps)f(ps)$ exceeds $y(p)$, the current social return per policeman. Let $p^*$ be the $p$ that maximizes (9); thus, if $p^*$ is positive, it is determined by the first-order condition

$$s(h - ps)f(ps) = \frac{\int_{0}^{ps} (h - b)f(b)\,\mathrm{d}b}{p}, \tag{11}$$

which for simplicity we will assume is uniquely satisfied.[11] The interpretation of $p^*$ is that it is the probability that maximizes social welfare per policeman; if police are used in a location so as to create the probability $p^*$ there, the payoff per policeman is highest. We can now state

[11] If there are multiple solutions to (9), Propositions 2 and 3 can be shown to hold with respect to the highest of these solutions. In the case of Proposition 2, the proof in the Appendix applies with slight modification (some of the inequality signs should no longer be strict).

**Proposition 2.** *If the available enforcement resources are such that $P \leq p^*$, then the optimal enforcement policy is to enforce with probability $p^*$ in one region of the city and not to enforce at all elsewhere. Specifically, $p(t) = p^*$ in $[0, P/p^*]$, and $p(t) = 0$ in $[P/p^*, 1]$.*

This result is proved in Appendix. The logic behind it is suggested by the following informal argument. Suppose that the social return per policeman is highest when the probability of enforcement is $p^*$, such as 30%. Then if one is assigning police resources, and there are not enough police to create a 30% likelihood in the whole city (that is $P < 30\%$), it is best to assign police so as to create a 30% probability in as large a region as one can. To think otherwise, for instance, to have a 25% probability in one region and a 5% probability elsewhere, cannot be as good, for by hypothesis when the probability is 30% the social return *per* policeman is highest; if police are spread so that they create 25% in one region and 5% elsewhere, the return per policeman in each area is lower by hypothesis, so those police raise social welfare less than social welfare is raised by having them concentrated so as to create a 30% probability in one region. Now when police are used to create a probability of $p^*$ in as much of the city as possible, the area in which this happens is $p^*/P$, for then the constraint (1) is obeyed: $(p^*/P)p^* + (1 - p^*/P)0 = P$.

It should be emphasized that Proposition 2 implies that, regardless of the underlying distribution of benefits $f(b)$), so long as $p^*$ is positive, it is *always* desirable to focus enforcement effort in one region of the city – and not to use any police elsewhere – when the available police resources are such that $P$ is less than or equal to the threshold $p^*$.[12]

Let us next discuss the situation where $P > p^*$, so that there are more enforcement resources than are needed to set $p^*$ throughout the city. In this case, as proved in the Appendix, there cannot be a region with zero enforcement. Hence, in view of Proposition 1, the problem to be solved is to maximize (6) over $p_1$, $p_2$, and $k$, subject to (5), where $p_1$ and $p_2$ are positive. The first order-condition for $p_1$ is:

$$w'(p_1) = w'(p_2) \tag{12}$$

or

$$(h - p_1 s)f(p_1 s) = (h - p_2 s)f(p_2 s). \tag{13}$$

The first order condition with respect to the size of the regions $k$ is:

$$w(p_2) - w(p_1) = w'(p_2)(p_1 - p_2). \tag{14}$$

These conditions determine the optimal solution, and it is quite possible that the solution is such that $p_1$ is unequal to $p_2$. The next result summarizes what we have just discussed.

**Proposition 3.** *If the available enforcement resources are such that $P > p^*$, then the optimal enforcement policy involves a positive probability everywhere. Either the optimal policy is uniform and equal to P, or, if the optimal policy involves two different probabilities, the first-order conditions determining the solution are:*

[12] Of course, $p^*$ depends on the distribution $f(b)$ (as well as on $h$).

$$w(p_2) - w(p_1) = w'(p_2)(p_2 - p_1) = w'(p_1)(p_2 - p_1), \quad (15)$$

where $kp_1 + (1 - k)p_2 = P$.

## 3. Discussion

We here discuss briefly a number of issues that bear on the model.

### 3.1. Knowledge of focused enforcement

For the focusing of enforcement effort to have an effect on deterrence, the probability of enforcement must, of course, be known by those to whom it applies. Thus, for instance, if enhanced traffic enforcement is undertaken by police on a highway, drivers must be apprised of this in order to be additionally deterred. The ability of enforcement authorities to communicate enforcement effort to different groups will vary according to context, but often it seems that enforcement authorities can accomplish this at little cost.[13]

### 3.2. Ability to escape focused enforcement

Presuming that individuals have knowledge of focused enforcement effort, it is obviously necessary that they cannot easily escape this enhanced effort for the effort to have an effect. If individuals could costlessly become members of the group subject to the lower level of enforcement effort, the focusing of enforcement effort would have no influence, and the situation would devolve into one in which a single likelihood of enforcement applies. However, individuals often would bear a cost in changing their group. If the group is one which travels along a particular route that is subject to greater enforcement, the cost would be that of taking an alternate route; this cost would often be non-trivial. The enforcement authority must, and often can, choose categories so that the ability of individuals to escape focused enforcement is limited.

### 3.3. Level of enforcement resources

We assumed in the analysis that there is a fixed level of available enforcement resources, and asked how optimally to employ these. But the level of resources, and notably, the number of police or other enforcement agents, can generally be increased, at a cost, or

[13] A point closely related to that of this paragraph concerns situations in which enforcement effort is in fact focused and individuals do not know where so, unless that is announced. For instance, there may be enhanced patrolling of some highway, or there may be cameras installed that take pictures of those who go through red lights at some intersection, but individuals might not know which highway is heavily controlled or which intersection has the cameras, unless this is announced. For the enforcement authority, not announcing the location of focused enforcement is equivalent to uniform enforcement, since individuals will (other things equal) assume the likelihood of focused enforcement is uniform in the city. Thus, in principle, the enforcement authority should announce where enforcement is focused only if that is socially desirable, along the lines that are described by the theory in this paper.

reduced, saving cost. If the cost of enforcement agents (or of other resources) is constant per agent, then a policy of focusing of enforcement effort cannot be strictly optimal:

Since the cost of hiring police will be the same in all regions, and since regions are identical, a policy which is optimal in one region must also be optimal in another. Hence, a uniform probability of enforcement in the city must then be optimal. However, for a variety of reasons, it will often be difficult for the number of enforcement agents to be increased at a constant cost, except in the long run; in the relatively short run, we may regard the number of police and other enforcement agents as somewhat capacity constrained. Among other things, police have to be trained, and also, once trained generally, have to learn about specific types of enforcement duties. Also, because of changing circumstances (an upswing in violations of a type, an increase in population, and the like), it will frequently be the case that the police in a given area or department find themselves constrained in numbers relative to their enforcement problem. Finally, there are political funding constraints that bear on the expansion of enforcement resources. For these various reasons, we would predict what we observe in fact, that the magnitude of enforcement resources often is not readily enlarged, opening up the possibility that focusing of enforcement effort might be beneficial.

### *3.4. Inequity of focused enforcement*

Because focused enforcement means that otherwise identically-situated individuals face different likelihoods of sanction, such an enforcement policy might be considered to be inequitable. However, when, as is usually the case, enforcement effort is allocated over many periods, concern about inequity can presumably be met by focusing enforcement effort first on one group and then on another, so as to achieve equal treatment of individuals over time.

## Acknowledgments

The authors thank Louis Kaplow and an anonymous referee for comments; Lando acknowledges the Center for Law, Economics and Financial Institutions at Copenhagen Business School for research assistance; and Shavell acknowledges the John M. Olin Center for Law, Economics, and Business at Harvard Law School for research assistance.

## Appendix

**Proof of Proposition 1.** We first demonstrate that the proposition holds for policies $p(t)$ that are presumed to be finite-valued. In this case, it suffices to show that any policy which has three different values on three regions can be replaced by another policy obeying the enforcement resource constraint which has only two different values on these regions, which is otherwise unchanged, and which produces at least as much social welfare. (For, by a finite number of such steps, any policy can be replaced by a policy obeying the enforcement resource constraint which has at most two different values and which produces at least as much social welfare as the initial policy.) Thus, assume that a policy $p(t)$ equals $p_A$ on region A, $p_B$ on region B, and $p_C$ on region C, where $p_A < p_B < p_C$. Observe that $p_B = \lambda p_A + (1 - \lambda)p_C$ for some $\lambda$ in (0, 1). Hence, we can expand the region B, where $p_B$ applies, by shrinking region A, where $p_A$ applies, and region C, where $p_C$ applies, in the proportions $\lambda$ and $1 - \lambda$, respectively, and the enforcement resource constraint will be maintained. We can also do the reverse, reduce the region B, by expanding the regions A and C in the proportions $\lambda$ and $1 - \lambda$ and the enforcement resource constraint will be maintained. Now one possibility is that expanding region B raises social welfare or leaves it unchanged (that is, $v(p_B) \geq \lambda v(p_A) + (1 - \lambda)v(p_C)$). If so, let us expand B until one of the regions, A or C, vanishes. Then, by construction, we will have only two probabilities in what had been A ∪ B ∪ C, the enforcement constraint will be obeyed, and social welfare will be at least as high as it had been. The other possibility is that reducing region B raises social welfare. If this is the case, we reduce B until it vanishes, and again, we will have only two probabilities in what had been A ∪ B ∪ C, the enforcement constraint will be obeyed, and social welfare will be higher. Next, we demonstrate that the proposition holds for any piecewise-continuous policy $p(t)$. Assume that the proposition is not true, so that social welfare $W^*$ under the optimal policy $p^*(t)$ exceeds social welfare $Q^*$ under the optimal policy among the class of policies which have at most two values. We wish to show that this leads to a contradiction. Consider a policy $p_n(t)$ that approximates $p^*(t)$ and that is constructed as follows: the interval [0, 1] is divided into $n$ mutually exclusive subintervals of length $1/n$; $p_n(t)$ is constant on each subinterval, and set equal in the subinterval to the mean of $p^*(t)$ in the subinterval. Hence, note that $p_n(t)$ satisfies the enforcement resource constraint. By definition of the integral, $\int_0^1 w(p^*(t))\,dt$ equals the limit of the integrals of $w(p_n(t))$ as $n \to \infty$ (for the $p_n(t)$ are step functions whose limit is $p^*(t)$). Accordingly, if $n$ is sufficiently large, the integral of $w(p_n(t))$ will be within any small $\varepsilon$ of $W^*$, and thus will exceed $Q^*$. But this involves a contradiction: the integral of $w(p_n(t))$ is less than or equal to the integral of some two-valued policy, since $w(p_n(t))$ is finite valued, and the proposition holds for finite valued policies; thus the integral of $w(p_n(t))$ cannot exceed $Q^*$. □

**Proof of Proposition 2.** Observe first that the social problem of maximizing (6) subject to (5) is equivalent to the problem of maximizing

$$kv(p_1) + (1 - k)v(p_2) \tag{A.1}$$

subject to (5), since $v(p) = w(p) - w(0)$ differs from $w(p)$ by a constant term.
Now suppose that the optimal policy is different from what is claimed. One possibility is that $p_1$ and $p_2$ are both positive, in which case (A1) can be rewritten as:

$$kp_1[v(p_1)/p_1] + (1 - k)[v(p_2)/p_2]. \tag{A.2}$$

Because either $p_1$ or $p_2$ must be different from $p^*$ by hypothesis, and because $p^*$ is the unique maximum of $v(p)/p$, (A2) must be less than

$$kp_1[v(p^*)/p^*] + (1 - k)p_2[v(p^*)/p^*] = P[v(p^*)/p^*] = (p/p^*)v(p^*) \tag{A.3}$$

which is the value of (Al) under the claimed optimal policy.

The other possibility is that $p_1$ alone is nonzero (if $p_2$ alone is nonzero, label it $p_1$). In that case (A.1) can be rewritten as:

$$kp_1[v(p_1)/p_1] \tag{A.4}$$

since $v(0) = 0$. Because $p_1$ is different from $p^*$ by hypothesis (A.4) must be less than

$$kp_1[v(p^*)/p^*] = P[v(p^*)/p^*] = (P/p^*)v(p^*). \tag{A.5}$$

(Note that $kp_1 = P$ since $p_2 = 0$.) Thus (A.4) is less than the value of (A.1) under the claimed optimal policy. □

**Proof of claim of Proposition 3.** When $P > p^*$ enforcement is everywhere positive: Assume otherwise, and note from Proposition 1 that an optimal policy which involves a probability of zero in a region can only involve one other probability, $p_2 > P$. Without loss of generality, we may assume the probability is $p_2$ on an interval [0, k] and the probability is 0 on [k, 1]. It is clear that we can select a small subinterval $J$ of [0, k] and a small subinterval $J$ of [k, 1] such that $p_2$ (length of $I$) $= p^*$ (length of I∪J), since $p_2 > p^*$. Hence, if the original policy is replaced by a policy in which the probability is $p^*$ in I∪J, social welfare will be higher, and the enforcement resource constraint will be maintained. This contradicts the supposed optimality of the original policy. □

## References

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

Polinsky, Mitchell, A., & Shavell, Steven. (2000). The economic theory of public enforcement of law. *Journal of Economic Literature*, *38*, 45–76.