Notes for Gary Becker’s “Crime and Punishment: An Economic Approach”

By Group 3 (Samuelson):
Traviss Cassidy, Antonios Koumpias, Jianxing Liang, and Yanxi Zhou

I. Introduction

- Key normative questions:
  - How many resources and how much punishment should be used to enforce different kinds of legislation?
- The optimal amount of enforcement depends on the cost of catching and convicting offenders, the nature of punishments, and the responses of offenders to changes in enforcement.
- Crime and punishment are approached as a constrained optimization problem to minimize social cost.

II. Basic Analysis

A. The Cost of Crime

- Legal notion: all violations that constitute a breach of governing rules such as felonies, traffic violations, tax evasion, embezzlement and other white-collar crimes.
- Political notion: “Crime has become a malignant enemy in America's midst. Since 1940 the crime rate in this country has doubled. It has increased five times as fast as our population since 1958. In dollars the cost of crime runs to tens of billions annually. The human costs are simply not measurable.”
  —Lyndon Johnson to the Congress of the United States, March 8 1965.
- Economic notion: the sum of Direct Costs of various crimes, Public Expenditures at the federal, state, and local levels on police and criminal courts, and Private Expenditures such as burglar alarms or guards. Private costs are vastly underestimated as official numbers do not incorporate personal behavioral responses such as riding taxis or living in the suburbs.¹

¹ Macro-extension on the linkages between crime and the economy: in a recent review of the literature, Steven Levitt concluded that a 5 percent increase in the unemployment rate generates about a 5 percent increase in property crime rates.
30 years prior to the paper, more tax legislation outnumbered criminal legislation.

In addition, the steadily declining ratios of currency (C) over all money (M), $\frac{C}{M}$, and currency over consumer expenditures ($X$), $\frac{C}{X}$, had spiked as well. More interestingly, a number of developments should have decreased $C$ ($\frac{dC}{dt} < 0$), such as the increased use of credit cards and further urbanization. Surprisingly, evidence suggests otherwise. The increase preference for cash is presumably associated with an increase in illegal business.

B. The Model

\[ H_i: \text{harm inflicted on society by offense } i \text{ (external diseconomy)} \]

\[ G_i: \text{gain to offender by offense } i \]

\[ D_i: \text{societal net cost} \]

\[ O_i: \text{number of offenses} \]

\[ H_i = H_i(O), \quad H_i' > 0, \quad H_i'' > 0 \text{ (increasing marginal harm to victims)} \]

\[ G_i = G_i(O), \quad G_i' > 0, \quad G_i'' < 0 \text{ (decreasing marginal gain to offenders)} \]

\[ D_i(O) = H_i(O) - G_i(O), D_i'(O) > 0 \quad \forall \, O > O_a \text{ if } D_i'(O_a) \geq 0, \quad D_i''(O) > 0 \]

- Cost of Apprehension and Conviction

\[ C: \text{Cost of enforcing the law} \]

\[ A: \text{level of “activity,” e.g., police officers, court personnel, specialized equipment} \]

\[ p: \text{probability an offense is cleared by conviction} \]

\[ C = C(A), \quad C' > 0 \]
\[ A \cong pO \]
\[ C_o = C'p > 0 \]
\[ C_p = \frac{\partial C(pO)}{\partial p} = C'O > 0 \]
\[ C_{pp} = C''O^2 > 0, h_{pp} = 0 \]
\[ C_{oo} = C''p^2 > 0, h_{oo} = 0 \]
\[ C_{po} = C_{op} = C''pO + C' > 0, h_{po} > 0 \]

- In the above model we assume that activity is measured only by chance of conviction and that \( \varepsilon_p = \varepsilon_o \). Relaxing the aforementioned, we may introduce \( a \), another explanatory variable of illegal activity, and devise a new, more general, State of the Arts function

\( a \): arrests and other determinants of illegal activity

\( A^{\text{new}} = h(p, O, a) \)

\( C^{\text{new}} = C(A^{\text{new}}) = C(p, O, a) \)

\( C_p > 0, C_o > 0, C_a > 0 \) as we may assume that \( h_p > 0, h_o > 0, h_a > 0 \)

Restrictions preventing corner solutions:

\[ C_{pp} = C''(h_p)^2 + C'h_{pp} \geq 0 \text{ unless } h_{pp} \text{ sufficiently negative (unlikely)} \]

\[ C_{oo} = C''(h_o)^2 + C'h_{oo} \geq 0 \text{ unless } h_{oo} \text{ sufficiently negative (unlikely)} \]

\[ C_{po} = C''h_o h_p + C'h_{po} \cong 0 \text{ unless } h_{po} \text{ sufficiently negative (unlikely)} \]

- Averaging out Cost per offense \( AC = \frac{C(p, O, a)}{O} \), Becker suggests a $500 figure. This is the lower bound on the marginal cost of crime \( (C_o > $500) \) such that \( C_{oo} > 0 \) holds throughout.

- Supply of Offenses

- Criterion for crime commitment: whether expected utility gained from the offense exceeds the utility the offender could achieve by using his time for other activities. Becker’s approach was at the time a departure from the sociological theories about the determinants of the supply of offenses.

\( E(U_j) \): offender’s expected utility from committing an offense
$Y_j$: offender’s monetary and psychic income from an offense

$$E(U_j) = p_j U_j (Y_j - f_j) + (1 - p_j) U_j (Y_j)$$

- In a cost-benefit analysis, every individual derives the number of offenses he will commit, taking into consideration three exogenous variables:

  $f_j$: punishment per offense

  $u_j$: other influences such as the $(p_j, f_j)$ of other crimes

  $$O_j = O_j (p_j, f_j, u_j), \quad O_{p,j} < 0, \quad O_{f,j} < 0, \quad O_{u,j} \text{ indeterminate}$$

- Implications for public policy:

  The social planner should not just determine the values of $p_j$ and $f_j$ but also should take under consideration the risk-aversion offenders exhibit. Risk-loving offenders are more responsive to changes in $p_j$ than in $f_j$. Statistically, offenders indeed tend to be more responsive to $p_j$ than to $f_j$, making them risk-lovers by expected utility standards. A risk-loving offender’s expected income from illegal (risky) activities is lower than his expected income from legal (less risky) activities.

Aggregating among offenders $j$, we may consider the market offense function:

$$O = O(p, f, u) \text{ where } p, f, u \text{ are the weighted averages of } p_j, f_j \text{ and } u_j$$

- Punishments

Forms: fines, probation, parole, jail time, capital punishment

---

### Figure 1: Number of persons on probation, parole or institutionalized for an average day in 1965 (in thousands)
Cost of Punishments to offenders

Fines reflect the crime’s monetary equivalent. However, there is no direct measure of the cost of institutionalization. An estimate would be the discounted sum of earnings foregone and the value placed on the restrictions in consumption and freedom. It will be greater the longer the prison sentence and the greater the potential earnings of the offender.

Cost of Punishments to society

Aside from collection costs, the cost inflicted by offenses cleared by fines is returned to society. However, imprisonment imposes costs on taxpayers (e.g., security personnel, food, etc.).

Total Societal Cost of Punishments, 

TCP = CP_{offenders} + CP_{society}: Total Societal Cost of Punishments,

TCP_{fines} = 0 as fines constitute an transfer of wealth from the offender to society

TCP_{imprisonment} > 0

Let \( b_i \) be the coefficient variable that maps punishment per type \( i \) of offense to social cost \( f' \) where \( i = \) fine, imprisonment, torture, etc.

Then, \( b_{fine} \approx 0 \) while \( b_{torture}, b_{probation}, b_{imprisonment}, \ldots > 1 \)

III. Optimality Conditions

- Goal: Choose values of \( p \) and \( f \) to minimize total social loss in real income from offenses, convictions, and punishments

L: The function that measures social loss

bpfO: total social loss from punishments

\( L = L(D, C, b_f, O) \)

- Assuming a loss function identical to the total social loss in real income, we formulate

OBJECTIVE FUNCTION: \( L = D(0) + C(p, O) + bpfO \)

\( \frac{\partial L}{\partial D} > 0, \frac{\partial L}{\partial C} > 0, \frac{\partial L}{\partial b_f} > 0 \)
First-Order Conditions:

\[
\frac{\partial L}{\partial \ell} = D'O_f + C'O_f + bpfO_f + b\ell O = 0
\]

\[
\frac{\partial L}{\partial p} = D'O_p + C'O_p + C_p + bpfO_p + b\ell O = 0
\]

- If \( O_f \) and \( O_p \) are not equal to zero, one can divide through by them.

\[
MC_f = D' + C' = -bpf\left(1 - \frac{1}{\epsilon_f}\right) = MR_f
\]

And

\[
MC_p = D' + C' + C_p \frac{1}{O_p} = -bpf\left(1 - \frac{1}{\epsilon_p}\right) = MR_p
\]

Where

\[
\epsilon_f = -\frac{f}{O_f} \quad \text{and} \quad \epsilon_p = -\frac{p}{O_p}
\]

As \( p \Rightarrow \) partly \( \downarrow C(O) \Rightarrow \) \( MC_f > MC_p \forall p, f \) policy makers may choose.
• Given that \( b > 0 \), lines intersect where \( MC = MR \), yielding the number of offenses that minimize social loss \( L \) where \( \varepsilon_p > \varepsilon_f \), the opposite of a profit maximizing firm as revenues are taking to be the number of offenses.

• Does Crime Pay?
  • Risk-loving offenders (assumption): \( \varepsilon_p > \varepsilon_f \)
  • Research by Smigel, 1965 and Erlich, 1967 indicates that \( \dot{p} \) and \( \dot{f} \) lie in regions that correspond to a preference for risk by offenders, suggesting that public policy has been consistent with the optimality analysis.
  • Consider the death penalty: it is rarely exercised, but the severity of the punishment at least partially offsets this low probability. However, judges and juries may be unwilling to convict offenders if punishments are set very high. This would fundamentally change \( C \)’s composition as the latter would also depend on \( f \):

\[
C = C(p, 0, \alpha) \text{ transforms into } C(p, f, 0, \alpha).
\]

• In earlier centuries when the means of apprehension were limited \( p \) was substantially lower. The legal system compensated for this low \( p \) by meting out draconian punishments:

\[
\uparrow f \text{ such that } D = D[O(p, f, u)] = \bar{D}
\]

IV. Shifts in the Behavioral Relations

• This section analyzes the effects of shifts in the basic behavioral relations—the damage, cost and supply-of-offenses functions—on the optimal values of \( p \) and \( f \).

1. The effect of a change in marginal damage
• An increase in the marginal damages from a given number of offenses, \(D'\), increases the marginal cost of changing offenses by a change in either \(p\) or \(f\). In this case, the optimal values of both \(p\) and \(f\) would increase, which would decrease the optimal number of offenses.
• Application: If the other components of the loss in income were the same, the optimal \(p\) and \(f\) would be greater for the more serious offenses. (Here, more serious offenses mean the offenses with more damage.)

2. The effects of change in cost

![Diagram]

- change in \(C'\)
  - The increase in the marginal cost of apprehension and conviction for a given number of offenses, \(C'\), has identical effects as an increase in marginal damages.
- change in \(C_p\)
  - An increase in \(C_p\) would increase the optimal number of offenses. The reason is that, the increase in \(C_p\) has no direct effect on the marginal cost of changing offenses with \(f\) and reduces the cost of changing offenses with \(p\), which reduces the optimal value of \(p\) and only partially compensates with an increase in \(f\).
- change in \(C'\) and \(C_p\)
  - An increase in both \(C'\) and \(C_p\) would increase the optimal \(f\), but can either increase or decrease the optimal \(p\) and optimal number of offenses, depending on the relative importance of the changes in \(C'\) and \(C_p\).
- Application:
  - An increase in the salaries of policemen increases both \(C'\) and \(C_p\).
  - Improved police technology or occupational ethics decrease both \(C'\) and \(C_p\). In this
case, the optimal f is reduced, though optimal p and optimal number of offenses may not be changed. This explains why the secular improvement in police technology and reform has gone hand in hand with a secular decline in punishments.

3. The effects of change in elasticity

- if b>0
  - An decrease in $\varepsilon_f$ increases $MR_f$, which increases the optimal number of offenses and decreases the optimal f that is partially compensated by an increase in the optimal p.
  - An decrease in $\varepsilon_p$ increases $MR_p$, which increases the optimal number of offenses and decreases the optimal p that is partially compensated by an increase in the optimal f.
  - An equal percentage of reduction in both $\varepsilon_f$ and $\varepsilon_p$ increases the optimal number of offenses and tend to reduce both p and f.
- if b=0
  - In this case, both marginal revenue functions lie along the horizontal axis. Therefore, changes in $\varepsilon_f$ and $\varepsilon_p$ have no effect on p and f.

• Applications:
  - Discriminate between offenses
  - Separate markets for offenses into different submarkets with different elasticities of supply of offenses. Higher prices would be charged in the submarkets with lower elasticities. Therefore, if b>0, loss would be lower by “charging” lower “prices,” that is, lower $p$s and $f$s in markets with lower elasticities.
  - Discriminate between people
  - Separate people that commits the same offense into groups with different responses toward punishments. Unpremeditated criminals are comparatively unresponsive to the size of punishments.
- if b changes
An increase in \(b\) increases the marginal revenue from changing the number of offenses by changing \(p\) or \(f\) and thereby increases the optimal number of offenses, reduces the optimal value of \(f\), and increases the optimal value of \(p\).

V. Fines

A. Welfare Theorems and Transferable Pricing

- Not only levels, but also slopes as incorporated into elasticities of supply (the right hand sides of equation (21) and (22)) could affect the optimization, because prices paid by consumers are not fully transferred to firms or governments.
- If \(b=0\), no social loss from punishments, then the elasticity of supply would drop out of equation (21); payment of the offenders is fully transferable.
- If \(b>0\), there is net social loss, then the elasticity of the supply determines the change in social costs caused by a change in punishments.
- For example, underdeveloped and Communist countries still use the draft, payment in kinds, and other non-transferable pricing punishments. The optimality conditions are significantly affected by the assumption about transferable pricing (\(b=0\) or \(b>0\)).

B. Optimality Conditions

- If \(b=0\), \(C=0\), equations (21) and (22) reduce to
  \[
  D'(\hat{O}) = H'(\hat{O}) - G'(\hat{O}) = 0
  \]
  which means net damage is zero and determines the optimal amount of offenses \(\hat{O}\). If marginal harm is always greater than marginal gain, the optimal amount of offenses should be zero. Since \(p\) does not affect \(C\), which is always zero in this case, \(p\) would be one, and penalties should be set sufficiently high to make \(O=0\).

- From the offender’s point of view,
  \[
  V = G'(\hat{O})
  \]  
where \(V\) is the monetary value of the marginal penalties. Then by equation (3) and (24),

  \[
  V = H'(\hat{O})
  \]

Since \(C=0\), \(p\) could be set to one without cost, then \(V = f\), and

  \[
  f = H'(\hat{O})
  \]
That is, a fine paid by the offender would fully compensate the “victims”—the rest of society.

- If \( b=0, \ C \neq 0, \) and \( p \) is still assumed to be 1, then

\[
D'(\tilde{O}) + C'(\tilde{O},1) = 0
\]  

(28)

Since \( C'>0, \) then \( D'<0, \) then the optimal \( \tilde{O} \) should be less than when \( D'(O) = 0, \) then

\[
f = G'(O) = H'(O) - D'(O) = H'(O) + C'(O,1)
\]  

(29)

That is, offenders have to compensate for the harm they do and the cost of catching them.

- If \( b=0, \ C \neq 0, \) and if \( f \) were fixed, and \( p \) were allowed to vary,

\[
D'(O) + C'(O, p) + C_p(O, p) \frac{1}{O_p} = 0
\]  

(30)

It is plausible that

\[
\frac{dC}{dp} = C', \frac{\partial O}{\partial p} + C_p > 0, \text{ then } C' + C_p \frac{1}{\partial O/\partial p} < 0, \text{ and } \\
D'(\tilde{O}) = -(C' + C_p \frac{1}{\partial O/\partial p}) > 0, \text{ therefore the optimal } \tilde{O} \text{ should be greater than when the costs are zero.}
\]

Whether we \( C \) increase or decrease \( O \) depends on whether penalties are changed through \( p \) or \( f. \)

C. The Case for Fines

- Social welfare is increased if fines are used whenever feasible.
  - Probation and institutionalization use up social resources.
  - A wise use of imprisonment and other punishments must know the marginal gains and harms, and in addition, the elasticities of response of offenses to changes in punishments.
  - Fines could restore the status quo ante. Other punishments could not compensate, and require “victims” to spend additional resources for punishments.
  - A fine can be considered the price of an offense, and so can any other form of punishment. But monetary units are preferred here as they are generally preferred in pricing and accounting.
  - If the goal is to minimize the social loss from offenses, and not to take vengeance or to inflict harm on offenders, fines should not depend directly on offender’s income.
    ➢ The implication of monetary value of a given fine being the same for all offenders: for offenders punishable by imprisonment, \( p \) is negatively related to income, because time is more valuable to rich offenders, and they tend to spend more money on
planning their offenses, on good lawyers, on legal appeals, etc. For offenders punishable by fines, \( p \) is positively related to income, because the time of poor offenders is comparatively less valuable, so they spend more time on how to avoid getting caught.

- Fines cannot be relied on exclusively whenever the harm exceeds the resources of offenders. In such cases, fines have to be supplemented with prison terms or other punishments. A flexible system of installment fines is suggested to enable offenders to pay fines more readily and thus avoid other punishments.
- That such flexible system is considered unfair by some people could be explained by imprisonment that places a low value on time in prison.

D. Compensation and the Criminal Law

- If the case for fines were accepted, the primary aim of all legal proceedings would become the same: not punishment or deterrence, but simply the assessment of the “harm” done by defendants.
- An action would be “criminal” precisely because it results in uncompensated “harm” to others.
- Antitrust is an example of the fundamental changes in the traditional approach to criminal law. If compensation were stressed, the main purpose of legal proceedings would be to levy fines equal to the harm inflicted on society by constraints of trade.
- Harm inflicted is not easily measured. However, civil suits still function reasonably well on the whole. As experience accumulates, the margin of error would decline, and rules of thumb would develop. Avoid irrelevant issues by focusing on the information most needed for intelligent social policy.

VI. Private Expenditures against Crime

- A variety of private actions include guards, accountants, locks and alarms, insurance coverage extended, parks and neighborhoods avoided, taxis used in place of walking or subways.
- Assuming each person tries to minimize his expected loss in income from crimes, optimal private decisions could be determined.

\[
L_j = H_j(O_j) + C_j(p_j, O_j, C, C_k) + b_j p_j f_j O_j
\]

\( H_j \): the harm to \( j \) from the \( O_j \) offenses committed against \( j \)
$C_j$: $j$’s cost of achieving a probability of conviction of $p_j$ for offenses against him, positively related to $O_j$, negatively related to $C$ (public expenditures on crime), negatively related to $C_k$, the set of private expenditures by other persons.

$b_j p_j f_j O_j$: expected loss to $j$ from punishment of offenders committing any of the $O_j$

- Usually $b$ is much greater than $b_j$, which is often less than or equal to zero. That is, most punishments result in a net loss to society, but they may produce a gain for the actual victims.
- Since $b_j$ and $f_j$ are primarily determined by public policy, $j$ is to choose $p_j$ to minimize $L_j$. Differentiating with respect to $p_j$, the optimality condition is

$$H_j + C_j + C_{jp_j} \frac{1}{\hat{O}_j \hat{p}_j} = -b_j p_j f_j \left(1 - \frac{1}{\hat{e}_{jp_j}}\right)$$

(32)

where $\hat{e}_{jp_j} = -\frac{p_j \hat{O}_j}{O_j \hat{p}_j}$.

- If $b_j < 0$, and if the left-hand side of (32), the marginal cost of changing $O_j$, were greater than zero, then (32) implies that $\hat{e}_{jp_j} > 1$, while the optimality condition given by (22) implies that $\hat{e}_p < 1$. The discrepancy could be explained by the fact that offenders can substitute among victims, therefore $\hat{e}_{jp_j}$ is probably much larger than $\hat{e}_p$.

VII. Some Applications

A. Optimal Benefits

- Model can be extended to analyze positive externalities (“benefits”)
- The analysis of advantages, benefits, and benefactors is almost perfectly symmetrical to the analysis of damages, offenses, and offenders

$B$: benefits from some benefactor

$p_1$: probability of “apprehending” benefactors

$a$: award per benefit

$\nu$: other determinants

$b_1$: fraction of $a$ that is a net loss to society

$A(B)$: net social advantages from $B$

$K(B, p_1)$: cost of apprehending and rewarding benefactors
$B(p_1, a, \nu)$: supply of benefits

$K', K_{p1} > 0$

\[
\frac{\partial B}{\partial p_1} \cdot \frac{\partial B}{\partial a} > 0
\]

\[\Pi: \text{profit function} \]

\[\Pi = A(B) - K(B, p_1) - b_1p_1aB \quad (33)\]

The goal is to maximize (33) with respect to $p_1$ and $a$. Our first-order conditions give us:

\[A'(B) \frac{\partial B}{\partial a} - K'(B, p_1) \frac{\partial B}{\partial a} - \left( b_1p_1B + b_1p_1a \frac{\partial B}{\partial a} \right) = 0 \]

\[A'(B) - K'(B, p_1) = \frac{b_1p_1B}{\partial B} + b_1p_1a \]

\[A'(B) - K'(B, p_1) = b_1p_1a \left( 1 + \frac{1}{e_a} \right) \]

Where $e_a = \frac{\partial B}{\partial a} * \frac{a}{B} > 0$ is the elasticity of benefits with respect to the per-benefit reward. Similarly, differentiating with respect to $p_1$ we get:

\[A'(B) \frac{\partial B}{\partial p_1} - K'(B, p_1) \frac{\partial B}{\partial p_1} - Kp_1(B, p_1) \left( b_1aB + b_1p_1a \frac{\partial B}{\partial p_1} \right) = 0 \]

\[A'(B) - K'(B, p_1) - Kp_1(B, p_1) \frac{\partial p_1}{\partial B} = b_1p_1a \left( 1 + \frac{1}{e_p} \right) \]

Where $e_p = \frac{\partial B}{\partial p_1} * \frac{p_1}{B} > 0$ is the elasticity of benefits with respect to the probability of “apprehension.” Note that if $b_1 > 0$, then $e_p > e_a$ because it must be that $b_1p_1a \left( 1 + \frac{1}{e_a} \right) > b_1p_1a \left( 1 + \frac{1}{e_p} \right)$.

- This means that the supply of good works by benefactors will be more responsive to the probability of apprehension than to the size of the award.
- This mirrors the fact that offenders tend to be more responsive to changes in $p$ rather than in $f$, which is shown to be associated with risk preference in footnote 19.
- However, for benefactors $e_p > e_a$ implies risk avoidance, as shown by the following:
EU: expected utility

\( Y: \text{income} \)

\[
EU = p_1 U(Y + a) + (1 - p_1)U(Y)
\]  \( (2') \)

Now, \( e_p > e_a \) if and only if

\[
\frac{\partial EU}{\partial p_1} \cdot \frac{p_1}{U} > \frac{\partial EU}{\partial a} \cdot \frac{a}{U}.
\]

Differentiating \((2')\) we get:

\[
\frac{[U(Y + a) - U(Y)] \cdot \frac{p_1}{U} > p_1 U'(Y + a) \cdot \frac{a}{U}}{\frac{[U(Y + a) - U(Y)]}{a} > U'(Y + a)}
\]

which holds if and only if \( U \) is concave. Thus benefactors are risk avoiders.

- While the optimal values of \( p \) and \( f \) would be in a region where “crime does not pay,” the optimal values of \( p_1 \) and \( a \) would be in a region where “benefits do pay.”
- The above analysis implies that the smaller the elasticities of response of inventors, the smaller should be the probability and magnitude of rewards.

B. The Effectiveness of Public Policy

- Becker defines “effectiveness” as the ratio of the maximum feasible increase in income to the increase if all offenses causing net damages were abolished by fiat.

\( O_o, p_o, f_o: \text{optimal values} \)

\( O_1: \text{offenses that would occur if } p = f = 0 \)

\( O_2: \text{value of } O \text{ that minimizes } D \)

Effectiveness is defined as:

\[
E = \frac{D(O_1) - [D(O_o) + C(p_o, O_o) + b p_o f_o O_o]}{D(O_1) - D(O_2)}
\]

- So effectiveness depends on two behavioral relations: the costs of apprehension and conviction and the elasticities of response of offenses to changes in \( p \) and \( f \).
- These elasticities are important—“crimes of passion” and crimes by young people are harder to deter than crimes by calculating adults—but effectiveness probably differs among offenses more due to differences in the costs of apprehension and conviction.
C. A Theory of Collusion

- The framework can be applied to the theory of collusion between firms.
- There are two costs of eliminating violations:
  - Discovering violations and “apprehending” violators
    - For a fixed $p$, this cost increases as the number of firms in the market increases
  - Punishing violators
- Because punishments like predatory price-cutting and physical violence hurt the collusion as a whole, firms in a collusion are assumed to choose probabilities of detection and punishments to minimize the loss from violations.

VIII. Summary and Concluding Remarks

- The main contribution of the paper is to demonstrate that optimal policies to combat illegal behavior are part of an optimal allocation of resources

Some empirical support for Becker’s model:

  - This was the first serious econometric investigation to find evidence of a pure deterrent effect of capital punishment.
  - The empirical analysis suggests that on the average the tradeoff between the execution of an offender and the lives of potential victims it might have saved was of the order of magnitude of 1 for 8 for the period 1933-67 in the United States.
  - Some later studies contradicted this finding, showing that Ehrlich’s results were sensitive to small changes in specification; however, as Gary Becker points out in *Uncommon Sense* (2009), the preponderance of the evidence is that capital punishment deters murderers.
  - Steven Levitt (JEP 2004) is nonetheless skeptical of the empirical evidence of a deterrent effect, noting that the probability of execution for a given murderer is less than one percent.

  - This paper uses the timing of mayoral and gubernatorial elections as an instrumental variable to identify a causal effect of police on crime.
  - Increases in the size of police forces are shown to be disproportionately concentrated in mayoral and gubernatorial election years.
- Increases in police are shown to substantially reduce violent crime, but have a smaller impact on property crime.