Economics of Crime
Introduction and models of criminal behavior

0.1 Basics of crime economics

- Definition: crime brings externalities, increases costs of actions
- Legal vs harm distinction. Economic approach:
  - Positive analysis: the choice of illegal behavior - valid in any contexts.
  - Joint underlying economics (punishment for crossing the line), de-emphasizing the legal details
  - Normative analysis: policy tailored to crime. Very different prescriptions for "victimless crimes".
  - Will use "crime" for all illegal behavior
- Efficient offense: would we want to eliminate all crimes even if it was costless?
- Rationality
  - strong vs weak
  - perfectly rational case
  - add "irrational" cost or benefits - impulse, morals, etc.
  - some offenders respond
  - individual vs aggregate - economic models mostly rely on aggregate, average responses
  - thinking about models of behavior: of course, both rational and non-rational elements are present
  - economics more powerful if the rational elements play a larger role
- Rational policy

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

- Fighting crime is expensive
- Many policy tools:
  - Deterrence: police, courts, prisons, fines
  - Reducing illegal gains
  - Increasing legal gains (labor policies, education)
  - Treatment
- Economics provides a framework for assessing the effectiveness.
- There are inevitable trade-offs. Some concern moral judgments and difficult-to-measure valuations, but some not.
- If the same crime rate (harm) can be achieved by a lower-cost policy, than that policy should be implemented.
- Questions:
  - Spend money on prisons or police? Necessary positive question: Do offenders respond more to the severity or probability of punishment?
  - Spend more on deterrence or on social crime-reduction programs?

The questions above are better answered if we know the magnitude of behavioral effects: By how much would crime change if a particular policy changes. We need empirical research to support our theories and show trade-offs.

1 Economics of crime

1.1 The economic approach to crime

Criminal law punishes wrongful acts such as murder, theft, burglary, etc. Why do we, as economists, bother about crime? At the first sight, many crimes, such as theft, seem to be a mere transfer of resources from the victim to the thief with no efficiency consequences. At the second sight, crimes have real costs:

- The offenders spend real resources on committing crimes and protecting themselves from being caught.
- Potential victims are protecting themselves against crime. Real resources are being spent on protection measures such as guns, locks, secure software, taxi rides (as opposed to public transportation) etc.
- Crime is essentially an externality. The offender imposes costs on some other party with which he is not in a contractual relationship.
- For a vast majority of crimes committed, the gain to the offender (G, I am starting to introduce notation that we will use throughout) is way, way lot smaller than the harm to the victim (H). Stolen property generally sells for a fraction of the price in the legal market. The value of life destroyed in a murder is rarely lower than the gain to the murderer from having someone dead. In principle, there could be such things as "efficient theft", "efficient rape", efficient murder" - crimes where $G > H$. We would actually want these crimes to happen. But in the world where we live in, our problem is not that we might be deterring a few efficient crimes; the problem is that we are failing to deter many inefficient crimes.
- For many crimes, there is a market substitute: if a thief values something more than the current owner ($G > H$), he can simply buy it from him rather than stealing. Voluntary exchange is a much cheaper way of reallocating resources than theft and other crimes.

To sum up, crime is a very costly activity. We would like to minimize the costs of crime. We try to do that by facing potential criminals with the possibility that they will be caught by the police, convicted by the courts, and punished by monetary fine, imprisonment, execution or some other way. The mechanism why through which the criminal law system reduces the costs of crime is **deterrence** - the risk of being caught and punished deters many people from committing crimes altogether, and it deters those who do commit some crimes from committing even more and harsher crimes. However, deterrence is not free - apprehending and punishing criminals consumes real costs such as police, judges, prison buildings etc. For that reason we do not want to eliminate crime altogether because the costs of deterring all offenses would not justify the benefit.

1.1.1 What do we mean by crime?

For a lawyer, crime is simply whatever the criminal statute defines as crime. An important distinction from other wrongful acts (torts) is criminal **intent**. An economic meaning of crime is slightly different:

First, there are so called **victim-less crimes** (prostitution, smuggling, drug dealing). These are voluntary exchanges that increase welfare, so for us economists they are a good thing, but they are criminalized in most jurisdictions. It remains a puzzle why. Arguments have been offered, but none of them is convincing. For example, these activities allegedly generate negative externalities. Then why do not we simply tax them? Also, some of these externalities (wars between drug gangs) are a direct consequence of the illegal status.

There are also minor violations of rules such as speeding or illegal parking that are being prosecuted and punished, yet we do not really think of them as crimes. Why? The socially optimal level of such offenses would be positive even if the enforcement were costless. For a (perhaps large) number of violators the gain is greater than the (expected) harm to victims. That is, could these violators negotiate with potential victims, they would agree that such violations do take place , as the Coase theorem would predict. (Posner gives this example: a hiker gets lost in the woods, is starving, and breaks into a cabin and eats some food.) For these violations, it is costly for the offender to refrain from the violation (if the hiker does not steal the food, he would die, a real cost). For "real" crimes, on the other hand, the criminal actually expends resources to commit the offense (a burglar wants to steal food from a cabin, so he buys the appropriate tools and spends time at night to get there and break in - a real cost).

1.2 Becker 1968 model

Becker (1968) is the seminal paper on economics of crime. He asked the following question: What is the socially optimal level of crime deterrence? Specifically, what are the values of p (the probability of catching and convicting criminals) and f (the punishment imposed on criminals when caught) that optimally deter criminals from committing crime? If deterrence were costless, the optimal number of offenses would be zero; we would set the probability of conviction and the penalty high enough so that all criminals were deterred. However, apprehension is costly (police, courts, collecting evidence, etc.) and so is punishment them (prisons). When we increase p or f, we must trade-off the benefits of reduced crime against the higher costs of apprehension or punishment.

1.2.1 Assumptions:

Let O be the number of offenses.

i) $H = H(O)$, $H' > 0$ be the total social harm from offenses, which is increasing in the number of offenses.

ii) $G = G(O)$, $G' > 0$ be the total gain to offenders, which is also increasing in the number of offenses.

iii) $D(O) = H(O) - G(O)$, $G < H \ \forall O$, $D' > 0$, is the net social harm, which is assumed to be always positive (so we are not assuming that some offenses are efficient) and increasing in the number of offenses.

iv) The costs apprehension (detecting, catching and convicting criminals, or put simply, the costs of police) is $C(p, O)$

where

$$\begin{aligned}\frac{\partial C}{\partial p} &= C_p > 0\\ \frac{\partial C}{\partial O} &= C'_O > 0\end{aligned}$$

That is, the cost of police activity is increasing in the probability of conviction (for a given number of offenses), and it is increasing in a number of offenses for a given level of probability. (One possible assumption, actually employed by Becker, $C(p, O) = C(pO)$, i.e., the total cost depends only on total number of offenders caught. However note that this assumption is too simplistic. It is certainly more costly to catch 100 criminals if there are 100 of them around than it is to catch 100 criminals if there are 5000 of them around.)

v) The criminals are rational and they have full information about the probability of being caught and the resulting penalty if caught. An individual j has a private supply function of offenses:

$$O_j = O_j(p_j, f_j, y_j)$$

where p is the probability of being caught and convicted, f is the punishment, and y is a vector of all other relevant variables, such as moral objections against crime, earning opportunities in the private sector, etc. For the law-abiding citizens, $O_j = 0$ for the realistic values of p and f.[1]

[1]Note on criminals' behavior: It is an old question in criminology as well as in economics whether the criminals are more easily deterred by a high punishment or a high probability of conviction. It turns out that the answer depends on whether the criminal is risk-averse or risk-loving. If the criminal is risk-averse, he is more deterred by a high punishment, even if it comes with a low probability, and vice versa.

Proof:

Let Y denote the wealth of the criminal if he commits a crime and is not caught (i.e., the income from legal activities plus the gain from crime G. The expected income is

$$EY = p(Y - f) + (1 - p)Y$$

Consider an increase in p accompanied by a simultaneous decrease in f such that the expected income remains unchanged. (This is a so called "compensated" increase in p. It is the same intellectual exercise as the Hicksian demand, where an increase in price is compensated by an increase in income.)

$$dEY = [(Y - f) - Y]dp - pdf = 0$$

Hence a compensated change in p requires

$$\frac{df}{dp} = -\frac{f}{p}$$

We want to see how a compensated increase in p changes the expected utility. If $dEU/dp < 0$, then an increase in the probability of conviction accompanied by a compensating decrease in the punishment decreases the expected utility of the criminal, and such a criminal is more easily deterred by a high probability of conviction rather than a high punishment:

$$\begin{aligned}EU &= pU(Y - f) + (1 - p)U(Y)\\ \frac{dEU}{dp} &\mid_{comp} = U(Y - f) - pU'(Y - f)\frac{df}{dp} - U(Y) = U(Y - f) - U(Y) + fU'(Y - f)\end{aligned}$$

Hence

$$\frac{dEU}{dp} \mid_{comp} \lesseqgtr 0 \text{ if } U'(Y - f) - \frac{U(Y) - U(Y - f)}{f} \lesseqgtr 0$$

A final note on the criminals: The model does not really assume that criminals are perfectly rational. All it assumes is that they do respond to incentives: If the gains from crime decrease (say because the contract law makes it harder to sell the stolen property) or the costs of crime increase (more police, harsher punishment, better opportunities in the legal sector), they will commit less crime. There is ample empirical evidence showing that criminals do respond to incentives.

vi) The aggregate supply of offenses is $O = O(p, f, Y)$

vii) Any punishment can be expressed in a monetary equivalent, so from the offender's perspective, we can treat punishments as fines. However, different forms of punishment have different social costs for a given cost to the offender. If the cost to the offender is f, the cost to society is bf, $b \geq 0$. For fines, $b = 0$, since fines represent a pure transfer from the criminal to the state without any social cost (realistically, b is somewhat greater than zero even for fines since there are some administrative costs of collecting the fines). For prisons, $b > 1$ since the prisoner suffers a loss f plus the state spends resources on prison buildings, guards etc.

viii) There is some exogenously given maximum possible level of punishment, $f_{\max}$. If the punishment takes the form of a fine, the criminals are able to pay $f_{\max}$.

1.2.2 Solving the model:

We do not want to convict and punish the criminals in order to "do justice" or to revenge on them or to keep them out of the street, but to minimize the social costs of crime, including the costs of conviction and punishment. Therefore the designer of the criminal justice system should solve the program

$$\min_{p,f} L = D(O) + C(p, O) + bpfO$$

We simply take the derivative with respect to p and f and set it equal to zero (check the original Becker's article for the second order conditions that guarantee that such a solution is indeed a minimum).

$$\begin{aligned} \frac{\partial L}{\partial f} &= D'O_f + C_O O_f + bpO + bfpO_f = 0 \\ \frac{\partial L}{\partial p} &= D'O_p + C_O O_p + C_p + bfO + bfpO_p = 0 \end{aligned}$$

which can be expressed in terms of elasticities

$$\begin{aligned} D' + C_O &= -bfp\left(1 - \frac{1}{\varepsilon_f}\right) \\ D' + C_O + \frac{C_p}{O_p} &= -bfp\left(1 - \frac{1}{\varepsilon_p}\right) \end{aligned}$$

The last expression is positive if the utility function is concave (the slope of the tangent at $U(Y-f)$ is greater than the slope of the chord), that is, when the individual is risk-averse. Therefore risk-averse individuals are more deterred by high punishment used with a low probability. The expression is negative when the utility function is convex, that is, when the individual is a risk-lover. Therefore risk-loving individuals are more deterred by a high probability of punishment. It is widely believed (and has been confirmed empirically, for example, in Block and Gerety (1996)) that criminals are risk-lovers.

where

$$\varepsilon_f = -\frac{\partial O}{\partial f}\frac{f}{O}$$

is (the absolute value of) the elasticity of offenses with respect to fines and

$$\varepsilon_p = -\frac{\partial O}{\partial p}\frac{p}{O}$$

is (the absolute value of) the elasticity of offenses with respect to the probability of conviction. We can interpret the left hand sides of equations (3) and (4) as the "marginal costs" of crime coming from the reduction in f and p, respectively. The marginal costs of increasing crime by a reduction in punishment (the LHS of eq. (3)) is always positive - reduced punishment induces more offenses, which yield higher direct social costs (D') and higher costs of apprehension and conviction for a given p ($C_O O_f$). A small reduction in p also induces more offenses, which yield higher direct social cost and higher costs of apprehension and conviction. On the other hand, a small reduction in p reduces the costs of apprehension and conviction for a given number of offenses ($C_p/O_p < 0$). The LHS of eq. (4) could in principle be negative if C_p were sufficiently high, but it will not be in the optimum: you want to go as far in reducing p until the marginal costs is positive and equal to the marginal revenue on the RHS. The signs on the LHS imply the following for the elasticities:

$$\begin{aligned}\left(1-\frac{1}{\varepsilon_f}\right) &< 0 \Rightarrow \varepsilon_f < 1\\ \left(1-\frac{1}{\varepsilon_p}\right) &< 0 \Rightarrow \varepsilon_p < 1\\ \varepsilon_p &> \varepsilon_f\end{aligned}$$

That is, we want to set the fines and probability such that the supply of offenses is inelastic. Also, we want to have the elasticity with respect to the probability of conviction to be higher than the elasticity with respect to the punishment. This implies that in an efficiently designed system, criminals are more sensitive to the probability of punishment rather than its severity. Compare this to the previous result that the criminals are more deterred by a high probability only if they are risk-lovers and that in fact the criminals are risk-lovers. The risk-loving attitude of criminals need not be something psychologically inherent, but rather an outcome of self-selection: A rationally designed system of law enforcement makes crime interesting only for risk-lovers.

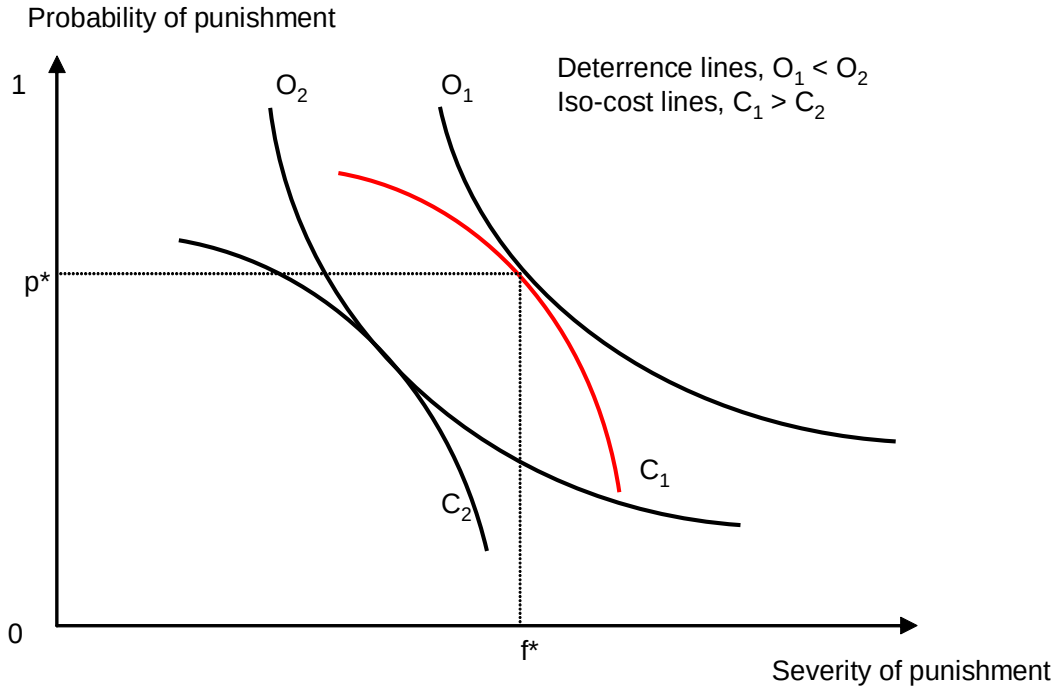
1.2.3 Graphical representation:

The comparative statics of the model are best demonstrated by graphically. Figures 1 and 2 decompose the problem of optimal deterrence into two stages: first, for any number of offenses that we choose to have, we select the values of p and f in order to minimize the costs of police and prisons. This gives us a "cost function", that is, the lowest possible expenditure on police and prisons needed to keep the number of offenses at a given level. Second, we select the number of offenses that we want to tolerate/eliminate.

Since criminals respond to both higher probability and severity of punishment, we can keep the number of offenses unchanged by increasing f and reducing p accordingly. This way, we can construct deterrence lines (sometimes called "iso-offense" lines), shown in Figure 1. Note that line O_1 represent a lower number of offenses than line O_2 since p and f are higher at line O_1. As for the costs, imagine there is a fixed budget to be allocated on police, courts and prisons. An increase in the resources spent on achieving a higher probability of punishment (i.e., spending more on police) has to be traded-off against a reduction in the severity of punishment (i.e., spending less on prisons) to keep the budget unchanged. This is shown as lines C_1 and C_2, where C_1 represents a bigger budget.

If we choose to have O_1 offenses, we should choose p^* and f^* as the least-cost way of achieving that number of offenses. If we repeat this exercise for every number of offenses, we obtain the total and marginal cost of deterrence associated with a given number of offenses. The marginal cost is depicted in Figure 2, which is drawn to show the marginal costs and benefits of eliminating the number of offenses by O_e, starting from some arbitrary number of offenses O_0.[2] For example, if we decided to eliminate offenses from O_2 to O_1, the marginal cost would be the difference between expenditure levels C_1 and C_2. The marginal benefit curve comes from the term $D'(O)$ in the model - the reduction in the net social harm coming from eliminating more offenses. The intersection of MSC and MSB curves gives the optimal number of offenses that we wish to eliminate. The figure is labeled so that the optimal of offenses eliminated, O_e^*, leads us to the level of offenses O_1. Then, the optimal choice of probability and severity of punishment are values p^* and f^* derived in figure 1.

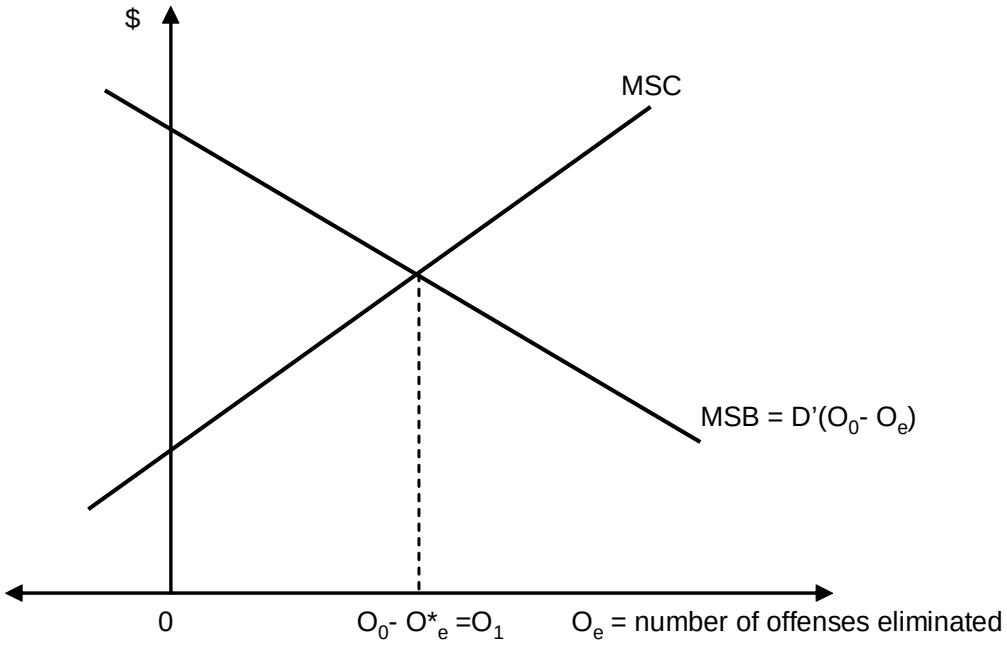
Figure 1: Trade-off between p and f



GRAPH: MC and MB of crime reduction

[2]Note that we can move into negative numbers on the x-axis, which would mean that instead of eliminating offenses, we choose to tolerate more offenses.

Figure 2: Marginal social costs and benefits of crime reduction



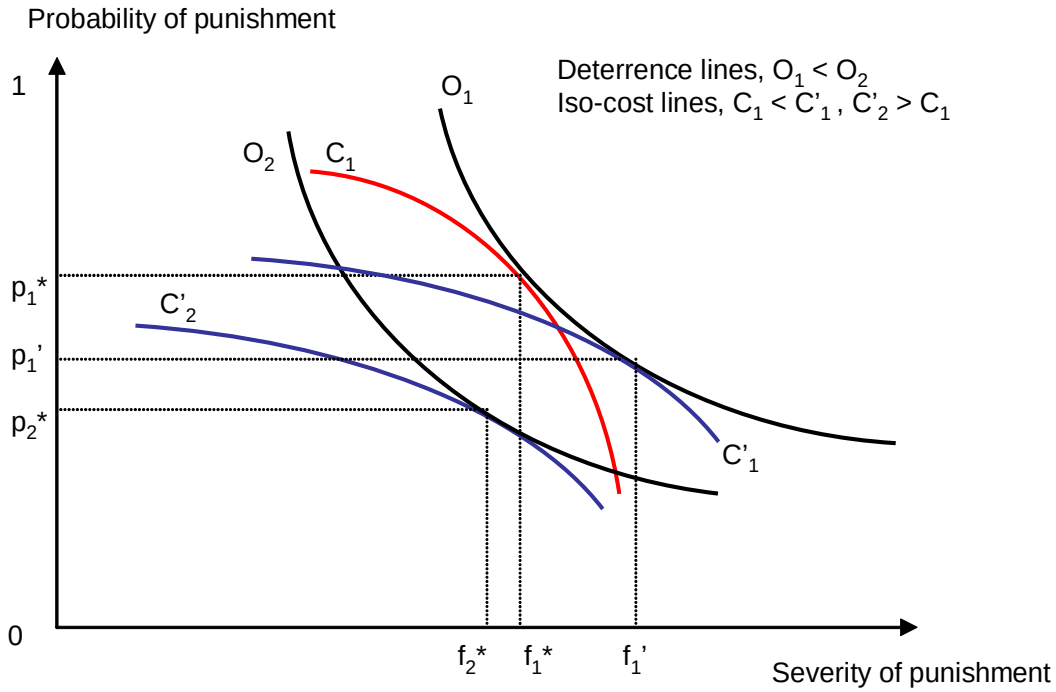
These graphs are very useful in analyzing the optimal response of the criminal justice system to changes in the variables of model, namely various shifts in costs and benefits of committing and deterring crimes. Consider the democratic revolution of 1989. Among other things, democracy greatly extended the rights of persons accused of committing crimes - police cannot benevolently arrest people, eavesdrop their telephone calls etc. If the police obtains evidence in an illegal manner, it is discarded in court, even though it shows the defendant is guilty. While these human rights reduce the risk that an innocent person is convicted, they also make it harder to convict those who are actually guilty. In the language of our model, the iso-cost lines shifts down and get steeper (Figure 3).[3] In order to keep the number of offenses at O_1, we need to spend more resources ($C_1' > C_1$), and holding the budget fixed, increasing p by 1 unit, requires sacrificing more in terms of a reduction in f. It then becomes optimal to reduce the probability of punishment and increase its severity to a point like (p_1', f_1') (this is equivalent to the substitution effect in consumer theory). However, we do not want to end up there - we are at the original number of offenses, while the costs of eliminating offenses has increased. In Figure 4, the MSC curve shifts up, implying that we want to eliminate fewer offenses (or, in other words, to tolerate more offenses, say O_2). Going back to Figure 3, this means we want to locate on the deterrence line associated with O_2 offenses, specifically at point (p_2^*, f_2^*), where the new iso-cost line C_2' touches the deterrence line O_2.

Observe that the optimal response of the criminal justice system to an extension of human rights involves a lower probability of conviction, higher number of offenses, and a higher expenditure on deterrence. (The model is ambiguous as to whether f will increase or not. The substitution effect moves it up, yet the willingness to tolerate more offenses moves it down.) The data is remarkably consistent with these predictions: After 1989, the Czech Republic (and other

[3]In what follows, we assume that the social harm from crime, or the response of criminals to p and f, remains unchanged.

post-communist countries) experienced a sharp decline in the probability of apprehension and conviction, a sharp increase in the number of offenses, an increase in police expenditures, and only a small decline in the length of prison sentences.[4]

Figure 3: Response to democracy: adjustment of p and f



1.2.4 Implications from Becker's model:

1. If deterrence were free ($C = 0$), the socially optimal number of offenses would be zero. We would just set p and f high enough to deter all offenses.

2. Since $C > 0, C' > 0$, the socially optimal number of offenses is positive.

3. For more harmful types of crime (those with higher D') both the probability and the punishment are higher.

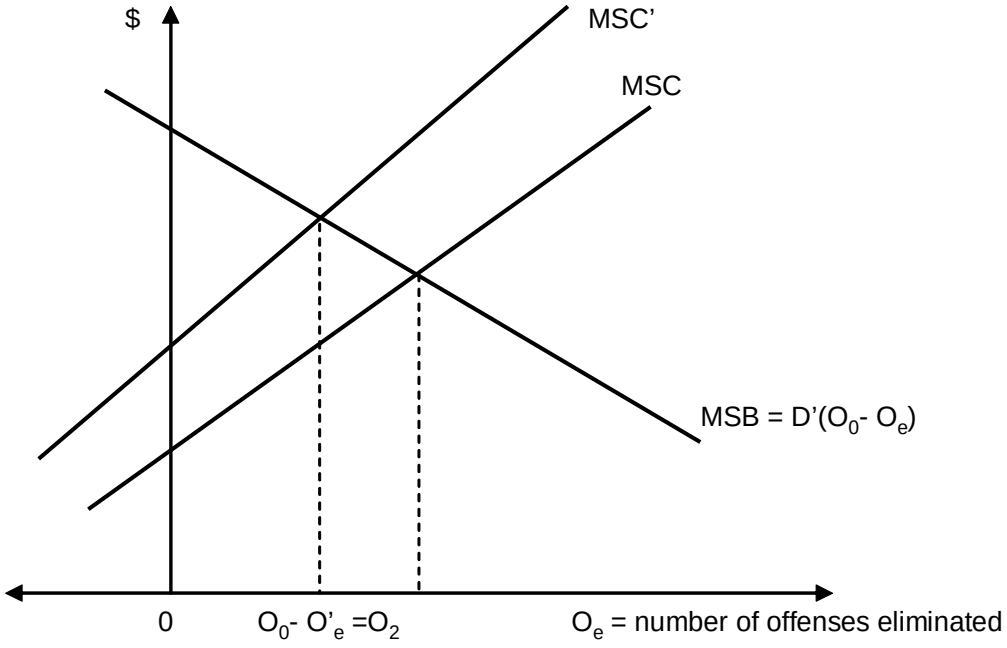
4. Exogenous reduction in p => higher f.

5. Fines are better than prisons. We achieve the same deterrence by setting the same f, regardless of whether we use prisons or fines. But since fines are socially costless, we save resources on prisons.

6. Assume that we use fines. Then for any desired level of offenses, one can minimize the social cost by setting the fine as high as possible ($f_{\max}$) while reducing the probability accordingly. This saves the costs of police, courts, etc. while achieving the same level of deterrence.

[4]Of course, many other factors were also behind the sharp increases in crime after 1989 - namely, the gains to offenders increased for many types of offenses.

Figure 4: Response to democracy: adjustment of crime



7. We can reduce crime not only by conviction and punishment, but also by improving the legal earnings opportunities of criminals, education, etc. - anything that changes the environmental variable Y.

1.2.5 Puzzles coming out of Becker's model:

We do not see the results 5 and 6 in the real world. Fines are used sparingly, while prisons predominate. We do not see harsh punishments with low probability, but rather the contrary.

Ignores the incapacitating effect of prisons: By keeping criminals off the street, they reduce the number of offenses that the criminals can commit (outside prison). Becker considers only the deterrent effect of imprisonment. There is no doubt that more prisons reduce crime, but it is very hard to distinguish empirically whether it is due to deterrence or incapacitation.

Fines are, in fact, expensive to collect because we do not perfectly observe the wealth of criminals. A person committing crime can "insure" himself against the risk of paying fine by nominally transferring his property to relatives etc. The fine may exceed the criminal's wealth.

Not much evidence on substitution between p and f (result 4). Lawyers don't tend to think this way (officially).

1.3 Extensions of Becker's model:

1.3.1 Why do we see mild punishments?

- Marginal deterrence (Stigler (1970)). If all crimes were punished by equal (and very high) punishment, then once you commit one crime, the marginal costs of committing other crimes is zero. So if the punishment is execution for both burglary and murder, nothing will stop you to kill a random witness of your burglary. The current punishment structure deters people from committing worse and worse crimes.

- Costs of error. It sometimes happens that people are convicted of offenses they did not commit. If the punishment for even the smallest violations were execution, people would take extreme precautions in order to avoid being a suspect. For example, if the speed limit is 50, everybody would drive below 40 in order to insure against the measurement error in the police radar. Plus, all efficient offenses would certainly be eliminated.

- Related point (Andreoni 1990) - also an implication of the cost of error and "reasonable doubt". In the civilized countries, courts attach very high cost on convicting an innocent person. Standard of proof: **beyond a reasonable doubt**. Reasonable doubt simply means small probability of making a type II error. As the length of prison term goes up, they really want to make sure they are convicting a guilty person $\Rightarrow$ some guilty guys are released $\Rightarrow$ p is a decreasing function of f. You can buy more deterrence by setting a moderate f and not going through much hassle in the courtroom. Extreme example: death penalty.

Q: Why are criminal attempts punished less severely than successfully completed crimes? E.g., if you shoot at someone and you miss, why do you get a lower sentence, even though your intention was obviously to kill? Quite the contrary, why punish you at all, since no harm was done?

1.3.2 Why prisons rather than fines?

- Incapacitation (see discussion above)

- Judgment-proof problem: A person whose wealth is less than the fine behaves as if the fine were only equal to his wealth, and so higher fines have no deterrent effect on him. Therefore, prison is a way how to impose a penalty that exceeds prisoner's wealth.

- In a way, we do not want the punishment to be too efficient. Fines are a transfer, so the criminal's loss is the government's gain. Fines would therefore give the government an incentive to prosecute and charge higher fines not just to get deterrence, but also to collect money (and be able to tax less). With imprisonment, the criminal loses and the government also loses (because it pays for the costs of prisons). Therefore, the government essentially "buys deterrence" and must weigh some costs and benefits (whether it faces the true costs is another question.)

1.3.3 Heterogeneity among offenders

- Why higher punishments for repeat offenders? This is a self-selection problem. If a person commits a crime, he signals that the current levels of punishment and probability were not sufficient for him (perhaps because he thinks he has a superior ability to escape from police). Therefore we impose a higher punishment for the second crime. Another reason is that prisons

effectively decrease the earnings ability of criminals. Therefore, after being released from prison, the ex-criminal has even stronger reason to participate in illegal rather than legal activities. Higher punishment for repeat offenses compensates for that.

- Rich versus poor offenders. There is a reason why to punish the rich with shorter prison sentences than the poor: Their value of time is higher, so they suffer more for a given prison term. To get the same deterrence, we should make the sentences declining in wages. On the other hand, there is a reason why to punish the rich more severely: They are better at releasing themselves from being convicted (can hire better lawyers), so they face a lower probability of conviction.

Reading list for this chapter

Introduction: Crime, rationality, the scope of criminal law

- Piehl, A. M., Williams, G. (2010). Institutional Requirements for Effective Imposition of Fines. In Controlling Crime: Strategies and Tradeoffs (pp. 95-121). University of Chicago Press.
- Donohue, J: Fighting Crime: An Economist?s View, Milken Institute 2005.
- Mungan, M. C. (2012). 2 The scope of criminal law. Research Handbook on the Economics of Criminal Law

Becker's model of crime and optimal punishment

- Becker, Gary S.: Crime and Punishment: An Economic Approach, Journal of Political Economy, Vol. 76 (2), 1968: 169-217.
- Cooter, R. and Ulen, T (2011): Law and Economics, 6th edition, Prentice Hall. Chapter 11

Refinements of the Becker's model

- Polinsky, A.M. and S. Shavell: The theory of public enforcement of law, in: Polinsky, A. M. and Steven Shavell (eds.): HLE, Chapter 6, 403-454.
- Ehrlich, I.: Crime, Punishment, and the Market for Offenses, Journal of Economic Perspectives 1996.
- Shavell, S: A Model of Optimal Incapacitation, American Economic Review, May 1987.
- Miceli, T: Deterrence, Incapacitation, and Repeat Offenders, working paper 2008

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

Národohospodářská fakulta VŠE v Praze