Economics of Crime
Criminal procedure

1 Criminal procedure

Criminal procedure is the algorithm (and rules) by which, when a crime happens, a suspect is identified and his guilt or innocence is determined, and sentence is set. Involves aspects such as:

- Arrest

- Miranda rights and rituals ("you have a right to remain silent...")

- Rights of the accused person

- Pretrial detention or bail

- Steps in prosecution, court hearings, rules of investigation

- Rules of evidence, standard of guilt

- Roles of attorneys

- Appeals

- Trial itself, judge and jury

- Prison and penitentiary system

Huge amount of literature is studying the prosecutors (their career objectives, choices to prosecute) - entirely U.S. based and hence heavily affected by the presence of plea bargaining. Plea bargaining has made its way to almost all criminal law systems in the world in different forms.

Let's focus more on "universal" features of enforcement.

1.1   Evidence and judicial errors

The purpose of the criminal procedure is to **release the innocent, convict the guilty**.

Information is imperfect - we never know for 100% that the guilty defendant is factually guilty. The criminal procedure guides the process of gathering, sorting, rejecting, evaluating information - the result should be a more precise indicator of the defendant's guilt or innocence. But errors are to some extent inevitable.

The way to think about this: The defendant is either factually guilty or factually innocent (G, I). The judge observes evidence E, and the evidence is indicative of guilt. Let $P(G|E)$ be the probability that the defendant is guilty, given the evidence. (That means, if we have enough defendants with the same evidence E, fraction $P(G|E)$ will be truly guilty. The judge needs to make a conviction / acquittal decision.

Null hypothesis: Innocent (legal default, innocent until proven guilty)

Two possible errors:

- Convict the innocent (Type-I error, rejecting a hypothesis that is true). If the person is convicted, the judge makes this error with probability $(1 - P(G|E))$.

- Acquit the guilty (Type-II error, failing to reject a hypothesis that is fals). If the person is released, the judge makes this error with probability $P(G|E)$.

Why are errors costly?

- Type II error: weaker deterrence

- Type I error: weaker deterrence (the true criminal is still out), ex-ante risk imposed on the innocence, mechanically more expensive because the society and the defendant bear the cost of punishment

- Intangible costs (injustice to a person)

Criminal procedure (or the various rules of the procedure) can be thought of accomplishing different objectives:

- Make the signal more precise (police investigation, trial, mandatory witnesses, rules against perjury)

- Reduce the Type-I error (right to counsel, right to remain silent, Miranda ritual, appeal).

- Reducing the abuse of police powers (restrictions on wiretapping, exclusion of illegal evidence) - this might increase Type II errors without reducing Type I.

Generally, measures that dilute information or only increase Type II errors should be avoided.

1.2 Andreoni (1991) model of reasonable doubt

This article is an effort to explain why we do not seem heavy punishments even though they would apparently deter crimes, especially the minor ones (why there is no death penalty for theft?)

Key insight: p **and** f **are not independent.** Judges/jurors respond to higher punishments by requiring higher standard of proof, effectively reducing the probability of conviction (conditional on apprehension). The resulting reduction in p may offset the deterrence effect of higher f (in fact, the sufficient condition for that is rather weak).

Legal background: the concept of reasonable doubt, why we in fact do not convict only if 100% sure.

The judge's/juror's decision making model assumptions:

1. The judge cares about the social cost/benefits of correct and incorrect verdicts.

2. f the cost of punishment to the defendant.

3. $c_1 =$ the social cost of convicting an innocent person $c_1 = \bar{c} + f$.

4. $c_2 =$ the social cost of incorrect acquittal; likely increasing with the severity of the crime, but not (necessarily) with f *given* the severity of the crime.

5. The cost of both types of correct verdicts normalized to zero.

6. $v(c_i) < 0$ ... the judge's disutility of an incorrect verdict; $v' < 0$, $v'' > 0$ (the judges are risk averse).

7. p ... the probability of the defendant's being guilty, as perceived by the judge after hearing the evidence.

Judge's decision making:

Given the perceived p, the expected utilities of conviction and acquittal:

$$
\begin{aligned}
EV^c &= p \cdot 0 + (1-p) \cdot v(c_1) \\
EV^a &= p \cdot v(c_2) + (1-p) \cdot 0
\end{aligned}
$$

Convicts if

$$
\begin{aligned}
EV^c &> EV^a \\
(1-p) \cdot v(c_1) &\geq p \cdot v(c_2)
\end{aligned}
$$

For any given c_1 and c_2, there is a critical level of p^* such that the judge decides to convict if $p \geq p^*$. This p^* in fact defines the "reasonable" doubt.

$$
\begin{aligned}
p^* &= \frac{v(c_1)}{v(c_1) + v(c_2)} \\
\frac{dp^*}{df} &= \frac{v'(c_1)[v(c_1) + v(c_2)] - v'(c_1) v(c_1)}{[v(c_1) + v(c_2)]^2}
\end{aligned}
$$

use the fact that

$$
\begin{aligned}
\frac{v(c_1) \cdot v(c_2)}{[v(c_1) + v(c_2)]^2} &= p^*(1-p^*) \\
\frac{dp^*}{df} &= \frac{v'(c_1)}{v(c_1)} p^*(1-p^*) > 0
\end{aligned}
$$

Now p^* is a policy parameter and p is a random draw from the distribution of possible $p's$ that may occur after the crime is committed. Then the probability of conviction is simply $\Pr(p \geq p^*)$. If p is uniformly distributed, the probability of conviction is $(1-p^*)$. Hence the prob of conviction falls as the penalty rises.

Criminal behavior:

1. b_1... The benefit of committing crime. The utility of not committing is normalized to zero

2. $b_2 = \bar{b} - f$... the net benefit of committing a crime and convicted

3. $u(b)$... criminal's utility, $u(b_1) > 0 > u(b_2)$, $u' > 0$, $u'' < 0$ (criminals are risk averse!) Weird assumption. What are the implications?

4. α ... the probability of apprehension $\rightarrow$ the probability of punishment if committing crime is $q = \alpha(1-p^*)$

Choice of committing the crime, defined by a critical value q^*; committ the crime if $q < q^*$.

$$
\begin{aligned}
EU &= (1-q) \cdot u(b_1) + q \cdot u(b_2) \\
q^* &= \frac{u(b_1)}{u(b_1) - u(b_2)} \\
\frac{dq^*}{df} &= \frac{u'(b_2)}{u(b_2)} q^* (1 - q^*) < 0
\end{aligned}
$$

(This does not take into account the adjustment of p to an increase in f.)

Implications of the changes in f

Can higher punishment encourage criminal behavior? That would happen if higher f increases the expected utility

$$
\frac{dEU}{df} = \frac{dq}{df} [u(b_2) - u(b_1)] - qu'(b_2)
$$

Substitute for

$$
\begin{aligned}
u'(b_2) &= \frac{dq^*}{df} q^* (1 - q^*) u(b_2) \\
q^* (1 - q^*) &= \frac{-u(b_1) u(b_2)}{[u(b_1) - u(b_2)]^2}
\end{aligned}
$$

Then

$$
\begin{aligned}
\frac{dEU}{df} &> 0 \text{ if and only if} \\
-\frac{dq}{df} \Big/ q &> -\frac{dq^*}{df} \Big/ q^*
\end{aligned}
$$

If the actual prob of conviction is greater (in percentage terms) falls by more than the prob of conviction sufficient to deter a criminal. Intuition: consider a marginal criminal for whom $q = q^*$. How plausible is this? Think that the EU is not increasing throughout: there is some f^* that minimizes the EU of the criminal. (So the expected utility is decreasing for $f < f^*$ but increasing for $f > f^*$). This then happens if the EU is convex in f at the minimum. Check the second derivative:

$$
\begin{aligned}
\frac{d^2 EU}{df^2} &> 0 \text{ iff} \\
-\frac{u''(b_2)}{u'(b_2)} &< -\frac{v''(c_1)}{v'(c_1)} + 2\frac{dp^*}{df} \frac{1}{p^* (1 - p^*)}
\end{aligned}
$$

The last term is positive, and the first two terms are the coefficients of absolute risk aversion of the criminal and the judge. The sufficient condition for this to hold is that the criminal has lower absolute risk aversion than the judge - seems very plausible.

Note: this prooves that there exists a range of f where the EU of the criminal is increasing; it does not say that it is true for the real world f. Just if we went too far with increasing the punishment we would hit that range.

Example: capital punishment. Not just this effect, but also the adverse effect of more expensive criminal proces (p^* is very high, it is also expensive to secure the evidence) on the resources available for deterrence of other crimes.

Finer implications: This effect should be weaker when the punishment is a fine instead of prison (for the same values of f).

General comment: example of an offsetting behavior. A policy is introduced with an alleged direct effect but its ultimate effect is mitigated through a variety of behavioral responses. This could spread to other parts of the system. Lawmakers cannot fully control p and f. (The same is true for f - judges have some discretion over the actual sentence).

The model widely accepted as theory, with surprisingly little empirical testing, and even less on judges. Positive prediction: p should fall as f rises. Also the magnitude of the effect is relevant.

1.2.1   Bjerk (2005) - Prosecutors and three strikes laws

Institutional features:

1. Three-strikes laws: 12 states, early-mid 1990's, details vary, but generally an offender with 2 prior violent felony convictions gets a life sentence on an another violent felony.

2. Felony vs misdemeanor, by selecting what crime would be the defendant charged, prosecutors are having extreme power.

3. Prosecutorial discretion: Prosecutors have a huge (near-unlimited) discretion to decide whether to actually prosecute a given suspect (hard to force them from outside to take action - the reasons why they do take action lie with career concerns/agency politics); decide the level of charges, plea bargaining tactics (you can plea guilty to a manslaughter. or go to trial for murder)

Data allows identifying an offender eligible for a three-strike sentence, if convicted of the charge for which he was arrested.

DiDiD set-up:

$$
\begin{aligned}
y_{icst} &= \beta_1 D(arrested \text{ for 3strike crime, post-law}) + \beta D(arrested \text{ 3strike crime}) \\
&\quad + \beta D(\text{arrested after 3 strikes law passed}) + \beta\,(\text{3strike state}) + \text{indiv. chars}
\end{aligned}
$$

Outcome of interest: probability that the charges are reduced to misdemeanor (then avoids 3-strike) Fairly strong result on β (higher prob by 8 percent). Little change in the characteristics of the population arrested for 3-strikes crime. Also regresses the likelihood of being convicted, if going to trial. No significant change, if anything, a statistically insignificant and small increase (by 3-4 percent). The selection issues are probably over-riding the behavioral response of the judges. Some explanations:

- Prosecutor reclassifies rather weaker cases to obtain some conviction, hence clearer cases go to trial as felonies.

- On the other hand defendants with a really weak case now dare to go to trial.

- Prosecutors can really think that the three-strikes are too harsh in individual cases, so they subvert the intention of the legislator and impose their view of justice (or just try to preserve the status quo).

- Strategic response by a prosecutor who actually seeks the maximal punishment: knowing that it may be very hard to obtain conviction at trial, rather drops some three-strikes cases, or reclassifies as misdemeanors, in order not to waste resources on these cases with very low chance of conviction.

1.3 Alternative criminal procedures

Grand trial is the "conventional" procedure. Every person should have a right to stand in the full trial. But it is extremely costly. In practice, simpler alternatives are used in many countries:

- **Penal order.** Judgment is proposed to the defendant by the court, he can choose to accept.

- **Plea bargaining.** Defendant bargains with the prosecutor about nature of the charge and the sentence. Deal has to be approved by the court.

- **Simplified/accelerated proceedings.** For certain cases with very easy fact finding or low proof standard, steps of the criminal proceeding can be skipped (administrative crimes).

Todays criminal systems are dependent on usage of this shortcuts and cannot grantee full and speedy trial to everyone.

A common fundamental economic logic in them:

- **Cheaper**

- **Convict with probability one or near-one**

- **Lower standard of evidence**

- **Resource-releasing hypothesis.**

  – Solving easy and menial crime by these procedures, resources should be released ("hand untied") to solve dangerous and resource intensive crimes (murders, robberies, etc.).

  – Contrary to previous point, it may motivate law enforcement to seek and solve crime using only above mentioned shortcuts to boost their crime statistics and disregard other, harder crimes.

There is a possible trade-off between deterrence and accuracy. The gap is lover if these alternative methods are sufficiently precise. Plea bargaining actually uses an additional information (the defendant's private info about guilt) and could lead to more precise sorting of the guilty and the innocent. The simpler procedure could reduce the errors among the serious cases.

Plea bargaining:

The prosecutor offers you that if you plead guilty, you get a shorter sentence than if you go to the trial and lose. Of course,if you win the trial you are free.

Superficial critique: **The guilty criminals are more likely to accept the bargain → they get lower sentence, less deterrence.**

Counterargument: **Plea bargaining frees up scarce resources to prosecute criminals.** Example from David Friedman: Assume the prosecutor has a fixed budget of $100,000, and there are 100 cases. If plea bargaining is not possible, he has to prosecute each case with $1000 per case → 50% of defendants go free. If he could get 90 defendants to plea guilty, he concentrates resources on the ten who go to court, and 90% of those are convicted. A defendant now faces a harder deal: accept a mild punishment or get convicted for a tough punishment with a 90% chance (without plea bargaining, he would face no choice and 50% chance of tougher punishment).

- If all criminals could "collude" and all plea non-guilty, they all would be better off.

- Plea bargaining increases the probability of being convicted (to which criminals are sensitive) at low cost.

- If the disparity between punishment at trial and plea are too large, the innocent would plea guilty.

Plea bargaining was recently introduced in Poland and Slovakia, works on a somewhat informal basis in Germany (everywhere limited to less serious crimes). Room for contribution to the U.S. literature - in the U.S., no way of testing what the world would look like without plea bargaining.

1.3.1 Lando (2005) - Size of the sanction should depend on the weight of evidence.

Not that just if evidence is weaker, we should be less willing to convict, but conditional on conviction, the sanction should be less severe if the evidence is weak.

Model: Takes the Polinsky and Shavell (2007) optimal deterrence model, adds doubt about the guilt - judges observe a weight of evidence x, higher x means more evidence. Exogenous threshold (1/2), but could be anything, to convict. Prob that evidence of strength x arises against the guilty perpetrator $f_g(x)$ and against the innocent $f_i(x)$. Conditional probability of guilt, given the evidence, is

$$P(g|x) = \frac{f_g(x)}{f_g(x) + f_i(x)}$$

This gives the probability of convicting a defendant, prob of convicting the guilty defendant, prob of convicting the innocent, and ex-ante prob that a criminal will be convicted. That in turn gives the supply of criminal acts. Potential criminals aware of the prob of drawing a given x, the corresponding prob of being convicted, and the sanction.

Intuition: Imagine you have two groups of cases, one relying on a super-strong evidence (DNA test), the other on somewhat weaker but still sufficient evidence (witnesses). Inevitably in both groups there are some people who are innocent but convicted, a higher fraction in the latter. You can punish with 4 years in prison in both types of cases or with 5 in the strong cases and 3 in the weaker cases. The social cost of punishment (imprisonment) is the same in both scenarios. But: In the strong group, there are more true criminals, so we're imposing a higher expected sentence - the higher sentence falls more often on the true criminals.

e.g. 1/2 of criminals charged, 1/2 of each type of evidence, 99% certain in the strong case, 90% in the weaker case.

Expected sentences:

4 years both: $0.5 \times ((0.5 \times .99 \times 4) + (0.5 \times 0.9 \times 4) = 0.4725 * 4 = 1.89$

5/3: $0.5 \times ((0.5 \times .99 \times 5) + (0.5 \times 0.9 \times 3) = 0.5 * (2.475 + 1.35) = 1.91$

Positive question: does it happen at all? Some anecdotal evidence. This should have implications for cases that are inevitably "cloudy" (corruption, rape).

Reading list for this chapter

- Andreoni, James: Reasonable doubt and the optimal magnitude of fines: should the punishment fit the crime? RAND Journal of Economics, Vol. 22 (3), 1991: 385-395.

- Sanchirico, Chris William: Evidence: Theoretical Models, chapter 9 in Procedural Law and Economics. Encyclopedia of Law and Economics, Second Edition, Edward Elgar 2012.

- Lando, H.: The Size of the Sanction Should Depend on the Weight of the Evidence, Review of Law and Economics, Vol. 1 (2), 2005.

- Leeson, P: Ordeals. Working Paper, 2010, (http://www.peterleeson.com/Ordeals.pdf)

- Bjerk, David: Making the crime fit the penalty: The role of prosecutorial discretion under mandatory minimum sentencing, Journal of Law and Economics, Vol. 48 (2), 2005: 591-625.

- Tsuchimoto, F. and Libor Dusek: Responses to more severe punishment in the courtroom: Evidence from Truth-in-Sentencing Laws, working paper available at http://home.cerge-ei.cz/ldusek/Papers/tispaperweb.pdf.

- Walsh, J. E. (2004). Tough for Whom? How Prosecutors and Judges Use Their Discretion to Promote Justice Under the California Three-Strikes Law. Crime and Justice Policy

- Garoupa, N. (2012). 9 The economics of prosecutors. Research Handbook on the Economics of Criminal Law, 231.

- Benson, B. L. (2010). The allocation of police. In Handbook on the economics of crime.+

- Mast, B. D., Bruce L. Benson, and David W. Rasmusen: Entrepreneurial Police and Drug Enforcement Policy, Public Choice, Vol. 104 (3/4), 2000: 285-308

- Dušek, L. (2014). Time to punishment: The effects of a shorter criminal procedure on crime rates. International Review of Law and Economics 2015.

- Dušek, L . and J. Montag (2016): The Effects of a Simpler Criminal Procedure: Evidence from One Million Czech Cases, working paper SSRN 2629116.

- Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. Proceedings of the National Academy of Sciences, 108(17), 6889-6892.

- Eren, Ozkan, and Naci Mocan. "Emotional Judges and Unlucky Juveniles, working paper 2016,

Národohospodářská fakulta VŠE v Praze