Excessive Sanctions for Governmental Misconduct in Criminal Cases

Richard A. Posner

Follow this and additional works at: http://chicagounbound.uchicago.edu/journal_articles

Part of the Law Commons

Recommended Citation

This Article is brought to you for free and open access by the Faculty Scholarship at Chicago Unbound. It has been accepted for inclusion in Journal Articles by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.
The fourth amendment, as is well known, forbids unreasonable searches and seizures by government officers; if the government tries to introduce evidence in a criminal trial that was seized in violation of the fourth amendment, the defendant can get the evidence suppressed. If the evidence is vital to conviction, this means that the defendant will be acquitted simply because the evidence was obtained illegally. This is the famous exclusionary rule of the law of search and seizure. The exclusionary rule is one example of a sanction for governmental misconduct: evidence that may be essential to convicting a dangerous criminal is suppressed to punish or to deter the government's violation of a law, here the fourth amendment.

Issues concerning sanctions for governmental misconduct may seem quintessentially legal, but they also have an economic dimension. I shall argue, building on an earlier paper in which I analyzed the fourth amendment's exclusionary rule, that economics can yield valuable insights into when sanctions for governmental misconduct are excessive.

These economic insights have two possible uses. One, of course, is to point the way to reform. The other, which is less obvious but from an
academic standpoint more interesting, is to explain the law as it is. The branch of legal scholarship with which I have been strongly identified has tried to explain the common law on the hypothesis that the law is designed to maximize economic efficiency.5

The hypothesis I shall explore in this article is that the common-law remedies for governmental misconduct in criminal cases are best explained by assuming that judges are preeminently concerned with economic efficiency, even though the underlying norms defining that misconduct are often not economic. My hypothesis does not pretend to explain how judges think about these cases; it is designed merely to explain the outcomes of a decisional process that judges usually rationalize in noneconomic terms. But just as economists consider it nonessential to the validity of their empirical results whether businessmen and consumers speak or think in the language of economics, so I believe it nonessential whether judges explicitly speak or think in that language.

I. THE TWO TYPES OF EXCESSIVE SANCTION

A sanction can be excessive from an economic standpoint in two ways. First, it can violate the Pareto criterion of efficiency6 by creating an avoidable deadweight loss. To explain the concept of "deadweight loss," I will use an example from the economics of sanctions for private (as distinct from governmental) misconduct: the choice between fines and imprisonment as the punishment for crime. To achieve a desired level of deterrence, society can, in principle at least, choose a fine that will be the exact equivalent of a term of imprisonment in the sense that the fine will impose the same private cost on the criminal. The social cost of the fine, however, will be smaller than the social cost of the equivalent term of imprisonment. The fine is just a transfer payment, whereas the term of imprisonment imposes deadweight losses—that is, losses not received as gains by anyone else—in the form of the criminal's forgone legitimate earnings and the costs of guarding him. Therefore, from an economic standpoint anyway, the fine is preferable.7

6. The Pareto optimum is a state in which no person can be benefited without a corresponding detriment to another person. If it is possible through a transaction to benefit one without detriment to any other, then the situation is not a Pareto optimum. See generally V. Pareto, Manual of Political Economy 103–80 (Schwier trans. 1971) (Pareto's theory of economic equilibrium).
Of course if the criminal is insolvent—as in the real world he often will be—this solution will not work. But if solvency is no problem, then a fine is Pareto-superior to imprisonment; that is simply the principle that I am interested in asserting here.

The second way in which a sanction may be economically excessive is that it may overdeter. The difference between this concern and the first can be seen most easily by imagining a choice between two fines, both collectible at zero cost from the defendant. If the smaller fine is set equal to the social cost of the defendant's crime, divided by the probability that he will be apprehended and convicted (which is simply to say, if the smaller fine is set at the optimal level), then the larger fine will be excessive. But it will be excessive not in the Paretoian sense of directly imposing an avoidable social cost—for I am treating the fine as a pure, costless, transfer payment—but in the sense of creating incentives for inefficient behavior.

Now in a world of perfect certainty, to be sure, fines could be infinitely large without imposing any social cost at all; the threat of having to pay the fine would deter anyone from engaging in the forbidden activity, and so the fine would never be imposed. But when the unrealistic assumption of perfect certainty is dropped, it becomes apparent that the threat of a very large fine (or of some other Draconian punishment) will induce people to avoid lawful behavior at the edge of the "forbidden zone" in order to minimize the probability of being falsely accused and convicted of the offense. The benefits of the lawful behavior that is avoided because of this risk are social opportunity costs of the excessive fine.8 Those costs provide the economic reason, or at least one important economic reason, for not imposing the death penalty on speeders; people would drive too slowly.

But it would be a mistake to conclude that a sanction really must be inefficient if it imposes an expected punishment cost greater than the formula for optimal punishment generates. The fine example makes it seem that all "excessive" fines are inefficient, but this is only because a fine is so easy to calibrate. But a fine is not always an available remedy, and the alternative sanction may not lend itself so easily to calibration as a sanction specified in dollars. In such a case it is necessary to compare the costs of overdeterrence with the costs of underdeterrence; if the latter costs are greater, the "excessive" sanction may not be excessive in a broader economic sense after all. The choice, then, is not between overdeterrence

---

and underdeterrence but between the optimal amount of deterrence and too much deterrence.

II. THE FOURTH AMENDMENT EXCLUSIONARY RULE

The exclusionary rule applied in fourth amendment cases illustrates both senses in which a sanction may be excessive from an economic standpoint. First, it violates the Pareto-superiority criterion because it imposes a deadweight loss—the suppression of socially valuable evidence—that would have been avoided if the misbehaving government official (the policeman who had made the illegal search) had been fined instead. Second, it produces overdeterrence because the private (and social) cost imposed on the government may greatly exceed the social cost of the misconduct.

To illustrate the latter, less obvious point, suppose that evidence that is indispensable to convicting a criminal is seized as an incident to some illegal search. Further suppose that the illegal search imposes a cost of $100 on the person searched in terms of lost time spent cleaning up after the searching officers, but that the loss to society from not being able to convict him can be valued at $10,000. If the probability of apprehending and convicting the illegal searcher is one, a fine of $100 would provide optimal (which is not to say 100 percent) deterrence of such illegal searches. The much larger "fine" that is actually imposed will overdeter, causing the government to steer too far clear of the amorphous boundaries of the fourth amendment compared to what it would do at the optimal fine level. The lawful searches that are forgone and the convictions of the guilty which those searches would have produced are social opportunity costs that the lower fine would have avoided.

These problems have long been recognized, and it is immaterial that they have not been formulated in explicitly economic terms. So we must ask why the exclusionary rule was ever adopted. The answer is consistent with economic analysis: the rule was adopted because until recently there was no alternative sanction for violations of the fourth amendment that did not cause severe underdeterrence. For reasons explained many years ago in an article by Caleb Foote, the natural (and superficially optimal) alternative remedy to the exclusionary rule—a tort action for damages against the government or its officers who engage in an illegal search—was for a long time unavailable. This unavailability was due to limitations ranging from lack of imagination by the courts in valuing intangible

---

10. Id.
losses in damage actions to broad concepts of official and sovereign immunity that usually made the judgments in these actions uncollectible.\textsuperscript{11}

Recent developments in tort law in general, and in the immunity doctrines regarding tort actions against governments and their officers in particular, have gone far toward solving these problems.\textsuperscript{12} The tort remedy has thus become a more practical alternative to the exclusionary rule. Admittedly, the tort remedy is not perfect. There remains in particular the problem of the loss that is large when aggregated over the large number of persons who may sustain it yet too small to give any one person an incentive to sue.\textsuperscript{13} An obvious remedy for this problem, though it has not yet been tried, is to set some minimum liquidated damage figure\textsuperscript{14} to which any plaintiff would be entitled (perhaps along with attorney's fees) in any case where liability is established. The minimum damages figure should optimally be set at the level that would induce just enough people to sue to make the total damages obtained equal to the total social costs inflicted by the police misconduct. To illustrate, suppose that the total costs of the misconduct to its victims is $10,000, that there are 1000 victims, each incurring an average cost of $10, and that if the minimum damages figure were set at $100, one in every ten victims would sue. Then 100 victims would sue, and the damages, $10,000 ($100 \times 100), would just equal the total social costs of the police misconduct.\textsuperscript{15}

Thus, the tort remedy for unlawful searches and seizures is now practical where once it was practically unavailable. As a remedy for police misconduct the tort remedy has the characteristics of the optimal fine that I discussed earlier.\textsuperscript{16} It is a transfer payment, and thus involves no (or more realistically, relatively little) deadweight loss. It also can be, and to some extent is, calibrated by judges and juries to yield the desired level of deterrence. As a result of the growing practical availability of the tort remedy, I would predict a shift away from the exclusionary rule.

In suggesting that economic analysis provides an explanation, and not just a criticism, of the law of sanctions for violating the fourth amendment, I derive additional support from an important traditional exception to the exclusionary rule: the refusal to bar prosecution on the basis of an

\begin{enumerate}
\item Id. at 496–504.
\item See Posner, supra note 4, at 64–68.
\item A general campaign of police harassment of some subgroup of the population would be an example of such a situation.
\item See Foote, supra note 9, at 496.
\item Of course, a drawback of this proposal is that it would give police officers no incentives to inflict costs of less than $100 on each victim of misconduct. Since the minimum penalty is $100, the officer might as well get his money's worth and do the full $100 of damage.
\item See note 4 and accompanying text supra.
\end{enumerate}
illegal arrest. A literal application of the exclusionary concept to the law of the fourth amendment would lead easily to a conclusion that if a person is arrested in violation of the fourth amendment he cannot be prosecuted. Take the case where someone is illegally arrested but not searched or even questioned, so that the arrest has no “fruits” of the kind that could be excluded from the prosecution’s evidence at trial. Suppose further that had the person not been arrested, he would not have been prosecuted. In this case, barring prosecution might seem to be the logical sanction for the illegal arrest, because it would correspond to suppressing the fruits of an illegal arrest. But barring prosecution would cause overdeterrence of an even more costly sort than the exclusionary rule involves.

Because the exclusionary rule only suppresses particular evidence, it does not prevent the prosecution from going forward if the prosecution has enough lawfully obtained evidence to convict the defendant. Barring prosecution altogether solely because the arrest was illegal would impose far greater social costs: not only the suppression of evidence that sometimes is (though often is not) indispensable to conviction, but also the dismissal of the charges in every case of illegal arrest. The law has stopped short of pushing the rationale of the exclusionary rule this far, and, as I have suggested, for a good economic reason.

I have argued that the overdeterrence problem that the exclusionary rule has created in search and seizure cases is solvable today because there is now a feasible tort alternative: a damage action against the misbehaving officer (or the government agency employing him) in which the court can nicely calibrate the damages to yield the optimal amount of deterrence. But in truth the tort approach has its own problem of overdeterrence. Police and other law-enforcement personnel are compensated on a salaried rather than piece-rate basis, so that even if they perform their duties with extraordinary zeal and effectiveness they do not receive financial rewards commensurate with their performance. At the same time, if their zeal leads them occasionally to violate a person’s constitutional rights, then the tort remedy will impose on these officers the full social costs of their error. There is thus an imbalance: zealous police officers bear the full social costs of their mistakes through the tort system but do not receive the full social benefits of their successes through the compensation system.

17. See, e.g., United States v. Crews, 445 U.S. 463, 474 (1980) (“An illegal arrest, without more, has never been viewed as a bar to subsequent prosecution ...”).

18. “Fruits” here refers to evidence that must be excluded under the “fruits of the poisonous tree” doctrine, because the evidence was obtained through an illegal search and seizure. See Nardone v. United States, 308 U.S. 338, 341 (1939).

19. For example, assume the government could not have discovered who or where he was but for some illegal dragnet that resulted in the illegal arrest.
We can fix this problem by immunizing police officers from tort liability, thereby externalizing some costs in order to eliminate a disincentive for the police to produce external benefits. But can we do this without also underdeterring police misconduct? We can—by ensuring that an officer’s immunity for misconduct (committed in good faith) is not extended to the agency employing him. This rule would essentially be one of respondeat superior without the employer’s usual right to indemnification: the agency would be fully liable though its employees would not. This rule would give the agency an incentive to prevent misconduct by its officers.20

III. EXCLUSION OF COERCED CONFESSIONS AND INVOLUNTARY GUILTY PLEAS

There are types of governmental misconduct where a tort remedy will not work and where, therefore, the exclusionary concept may be optimal despite its inherent overdeterrence. A good example is the coerced confession or involuntary guilty plea, extracted in violation of the fifth amendment’s self-incrimination or due process clause.21

Suppose that a criminal defendant could not exclude a coerced confession from evidence in his criminal trial but could (as, in principle at least, he can) bring a tort action seeking damages for the violation of his rights. Consider two distinct types of tort case. In the first, the criminal defendant proves that the confession is unreliable because it was coerced, and that the other evidence was insufficient to sustain his conviction. In such a case the appropriate measure of his tort damages would be the costs to him of whatever punishment had been imposed. But obviously the cheaper and more efficacious remedy in this case is simply to bar the use of the coerced confession from evidence at his criminal trial. That obviates the punishment, and so avoids the difficult and uncertain task of measuring the costs of punishment to the unjustly imprisoned defendant.

This case differs from a search and seizure case in that the latter involves no issue of reliability of evidence and hence no issue of guilt. In the search and seizure context, the private costs of punishment are not

20. See Posner, supra note 4, at 64–68, for a fuller discussion.
21. The fifth amendment provides in part: “No person . . . shall be compelled in any criminal case to be a witness against himself, nor be deprived of life, liberty, or property, without due process of law . . . .” U.S. CONST. amend. V, cl. 3.

Purists would limit the self-incrimination clause to in-court statements, thus excluding confessions coerced by the police before trial though introduced into evidence at trial. But that would not alter my analysis here.

Because my analysis of involuntary guilty pleas is symmetrical to that of coerced confessions, I will confine my discussion to the latter.
equal to, and usually are much greater than, the social costs of the government's misconduct. Where the defendant is not in fact guilty of a crime, the private and social costs of punishment coincide and become the proper measure of damages; more simply, this convergence provides a reason for excluding the evidence from trial.

The harder case is where the confession, although coerced, is reliable evidence. Perhaps the confession was corroborated, or perhaps the coercion was not so great as to have made the defendant confess unless he really was guilty. In this case, suppressing the confession at trial would overdeter just as much as would excluding evidence obtained by an illegal search.

There are three possible responses to this problem:

Option 1: Limit the substantive right under the fifth amendment to cases where the presence of coercion throws a serious doubt on the reliability of the confession. Then no violation would occur when the confession was corroborated or when the coercion was too mild to create a substantial doubt that the confession was true.

Option 2: Limit the tort remedy to the defendant's lawful interests, narrowly defined, as in the search and seizure case. Hence, if the police used the "third degree" to extract a confession that was corroborated or otherwise validated, the defendant could not suppress the confession in his criminal trial; but he could obtain damages for deprivation of food or sleep or for any other injury to his incontrovertibly lawful interests that resulted from the third-degree methods used to extract the confession.

Option 3: Apply the exclusionary rule. This, of course, is the current approach of the law.

Option 1 exceeds the scope of this paper; it probably also exceeds the realistic scope of judicial authority to reexamine settled doctrines. It may well be that the fundamental concern underlying the fifth amendment is with the reliability of the guilt-determining process; that the Framers thought an inquisitorial system of justice less reliable than an adversary one. But the fifth amendment also reflects a view that a person should not be forced to assist in his own conviction even if he is guilty.

The origin of this view is in Hobbes, who argued that because a person agrees to subject himself to the authority of the state only for the sake of self-preservation, the agreement lapses if the state tries to take away his life.

22. See United States v. Yurasovich, 580 F.2d 1212, 1215 (3rd Cir. 1978).
23. E.g., Tehan v. United States ex rel. Shott, 382 U.S. 406, 415 (1966) (privilege against self-incrimination premised on the philosophy that even the guilty should not be convicted unless the prosecution carries the burden of proving guilt).
24. Hobbes wrote: "If a man be interrogated by the sovereign, or his authority, concerning a
If coerced confessions were particularly effective in reducing crime and thereby increasing personal security *ex ante*, we might say that by joining in a civilized society a person really does waive any right to refuse to cooperate in his own destruction.

But all this is merely to concede that the underlying norms which define governmental misconduct in criminal cases are not merely economic norms. This places a constraint on remedial policy. Concretely, it forecloses Option 2. If a conviction that rests upon a coerced confession is unjust even if the defendant is clearly guilty, then no relief will be adequate that does not undo the effects of the conviction.

Some think a conviction based on illegally seized evidence also unjust despite the fact that there usually is no question about the reliability of such evidence. If they are right, the exclusionary rule cannot fairly be criticized as producing overdeterrence. But I assume that they are wrong—that the objection to unreasonable searches and seizures is not that they render criminal proceedings which use their fruits unfair, but that they invade collateral interests in property and tranquility which now can be fully protected by tort remedies.

The coerced confession and involuntary guilty plea are of course only two examples of procedural concepts designed to ensure the reliability and fairness of the criminal justice system. The same analysis that led me to conclude that exclusion is the natural and not the overdeterrent remedy for violations of the rules forbidding coerced confessions and involuntary guilty pleas compels a similar conclusion in the case of other rules—rules allowing the criminal defendant to be present at his trial, to have counsel to assist him in his defense, to be allowed to confront the witnesses against him, and so on.

But in all of these areas, including the coerced confession and involuntary guilty plea problems with which I began, there is a principle limiting the use of exclusion as a remedy. That principle, the doctrine of harmless error, is the last subject I shall discuss in this paper.

IV. HARMLESS ERROR

Suppose that the judge in a criminal trial admitted, in violation of the rules of evidence, certain hearsay evidence damaging to the defendant, but that so much admissible evidence of guilt was presented at the trial that the probability was very slight that excluding the hearsay evidence would have led to the defendant’s being acquitted. In such a case the ap-

---

crime done by himself, he is not bound, without assurance of pardon to confess it; because no man . . . can be obliged by covenant to accuse himself." T. Hobbes, Leviathan, pt. 2, ch. 21 (Collier ed. 1962).
pellate court would invoke the doctrine of harmless error to uphold the conviction notwithstanding the trial judge's error.\textsuperscript{25}

A superficially attractive justification for this result within the economic framework of this paper is that if the outcome of a retrial really is foreordained, then reversing the conviction will have a single Pareto-inferior consequence: to impose a deadweight loss consisting of the expenses of the retrial, a pure debit on the social books. But this analysis is artificial, for if the result of a retrial really is foreordained, then the criminal defendant's only possible incentive for seeking a retrial is to give the prosecutor an incentive to plea bargain by delaying the prisoner's incarceration and imposing costs on the prosecutor. And because settlement is always cheaper than litigation when the outcome of litigation is known with certainty, all of these cases will be settled; none will be retried.

The more interesting case, and no doubt the empirically more important one, is where there is some residue of uncertainty whether the defendant will be convicted when retried. Here retrial is more than a hypothetical possibility, because if the prosecution and defense cannot converge in their estimates of the probability of conviction on retrial they may find it cheaper in an expected-value sense to litigate than to settle.

How can the cost of retrial be a deadweight loss when by hypothesis some probability exists that this time the defendant will be acquitted? The answer lies in the fact that the probability that the defendant will be acquitted on retrial exceeds the probability that he is innocent. The reason for this divergence between the two probabilities is not the heavy burden of proof in a criminal case, for no one would propose to invoke the harmless-error rule unless the evidence that remained after setting aside the improperly admitted evidence proved the defendant's guilt beyond a reasonable doubt. The reason rather is that a nontrivial possibility always exists that a jury (or even a judge) will acquit a guilty person—because the jury either does not like the law,\textsuperscript{26} does not understand it, or cannot apply it correctly to the facts of the case.

The power of the jury to acquit is absolute; our system of laws contains no such thing as a motion for a directed verdict by the prosecutor. But the right of the jury to acquit is more limited, for the jury has only the power—and never the right—to acquit on grounds that the law does not recognize.\textsuperscript{27} Therefore, there is no paradox in stating that the acquittal of a guilty person imposes social costs, even though the acquittal cannot be corrected by appeal to the trial judge or to a higher court.

\textsuperscript{25} See \textit{Fed. R. Crim. P. 52}.


\textsuperscript{27} This distinction is stressed in P. DEVLIN, THE JUDGE 117–48 (1979).
All this is prefatory to interpreting the significance of the harmless-error rule. Its significance is that if the evidence which remains in the case after all erroneously admitted evidence is laid aside shows conclusively that the defendant is guilty notwithstanding what an errant judge or jury might conclude on retrial, the conviction will be affirmed. The reason, I think, has to be the concern with overdeterrence. If a person is guilty beyond a reasonable doubt on the basis of evidence both reliable and just, then a retrial will impose either a deadweight loss in the form of litigation expenses that will not change the outcome of the first trial, or an equally or (probably) more serious social cost resulting from the acquittal of a guilty person and consequent reduction in the deterrent and incapacitative effects of criminal punishment. These costs are excessive relative to the governmental misconduct, which by definition is slight since the defendant would in all probability have been convicted anyway.

What the harmless-error rule does, then, is to identify a type of governmental misconduct whose social costs are much lower than the social costs of attempting to deter the misconduct by overturning the conviction and forcing a retrial. This functional definition has the value of guiding the rule's application more dependably than by chewing over the connotations of the word "harmless." The functional approach implies, for example, that the rule should be interpreted more liberally for grave crimes. The graver the crime, the more the parties are likely to invest in the litigation process itself, and that greater investment should increase the accuracy of the guilt-determining process. Thus, if the trial court concludes that the untainted evidence proves the defendant's guilt beyond a reasonable doubt, this judgment is likely to be more reliable in a case of serious crime than in a trivial one. The appellate court's finding of harmless error will thus be more firmly based in such a case.

Therefore appellate courts can be expected to hold more errors harmless in grave rather than in trivial crimes. This counterintuitive implication follows directly from what I have called the functional, which is to say the economic, approach to analyzing the harmless-error rule, and it is an empirically testable implication of that approach.28

---

28. An initial effort to test the implication has not, however, been very successful. A random sample of recent federal court of appeals criminal decisions involving harmless-error issues yielded 42 usable observations, in 25 of which the harmless-error doctrine was applied, and in the remaining 17 of which the error was deemed reversible. Using length of sentence as the measure of the severity of the cases, I found that the average length of sentence in the first group was 11.36 years and in the second 12.42—which is contrary to my hypothesis. Details of the study are available from the author, and are on file with the Washington Law Review.
V. CONCLUSION

In devising remedies for governmental misconduct in criminal cases, courts, I have argued, have been guided by concerns articulable, if rarely articulated, in terms of economic efficiency. The underlying substantive concepts that define governmental misconduct need not be economic, and in the case of rules protecting the fifth amendment’s self-incrimination clause seem not to be economic at all. But the remedial scheme that courts have created seems responsive to efficiency concerns. If so, this is further evidence that economics has profoundly influenced the structure of the law.