Yale Law School
Yale Law School Legal Scholarship Repository

John M. Olin Center for Studies in Law, Economics, and Public Policy Working Papers

Yale Law School Other Scholarship

12-11-2002

The Theory of Value Dilemma: A Critique of the Economic Analysis of Criminal Law

Dan M. Kahan
Yale Law School

Follow this and additional works at: http://digitalcommons.law.yale.edu/lepp_papers

Recommended Citation
Kahan, Dan M., "The Theory of Value Dilemma: A Critique of the Economic Analysis of Criminal Law" (2002). *John M. Olin Center for Studies in Law, Economics, and Public Policy Working Papers.* Paper 280.
http://digitalcommons.law.yale.edu/lepp_papers/280

This Article is brought to you for free and open access by the Yale Law School Other Scholarship at Yale Law School Legal Scholarship Repository. It has been accepted for inclusion in John M. Olin Center for Studies in Law, Economics, and Public Policy Working Papers by an authorized administrator of Yale Law School Legal Scholarship Repository. For more information, please contact julian.aiken@yale.edu.

The Theory of Value Dilemma: A Critique of the Economic Analysis of Criminal Law

Dan M. Kahan[*]

Criminal law can justly lay claim to being the native domain of law and economics. From Bentham to Becker, no area of law has been more brightly illuminated by the radiance of economic logic.

Not coincidentally, in no other area of the law has the conflict between economics and ordinary moral sensibilities been perceived to be so sharp. When we are deciding how to apportion punishment, our instinct is to look backwards: wrongdoers should be condemned in proportion to the reprehensibility of their acts. I'll call the position that punishment should flow directly from these intuitions *intrinsic value retributivism*. Economics, in contrast, uses the forward-looking idiom of deterrence: punishment is justified if and to the extent that it produces a desired state of affairs; if the optimal degree of punishment so derived is out synch with intuitions about the reprehensibility of a particular crime, so much the worse for those intuitions.

My goal in this essay is to demonstrate that economic analysis lacks the conceptual resources needed to criticize intrinsic value retributivism in this fashion. The economic theory necessarily presupposes some theory of value that tells us what states of affairs we are trying to maximize through punishment. To construct that theory, the members of a community must consult their shared intuitions. Once a theory of value so derived is plugged into deterrence theory, however, the results won't differ in any material respect from the ones the law would produce if informed explicitly by intrinsic value retributivism.

This claim, which I'll call the *theory of value dilemma*, can be reduced to three sub-claims. The first is that the economic analysis of criminal law depends on an externally specified theory of value; without some account of what state of affairs the law is trying to maximize, the consequentialist logic of deterrence can't get off the ground. Second, this theory of value is essentially political; that is,

---

[*]Professor, Yale Law School.

there's nothing internal to economic approach that justifies discounting the valuations implicit in the decisions of a democratic political community's legal institutions, and hence nothing internal to economics that forecloses political advocacy of any particular theory of value. And the third and final sub-claim is that these politically derived evaluations will, as a practical and conceptual matter, always dominate the distinctive normative components of the economic theory of deterrence. I'll discuss each of these points in turn.

1. The Economic Theory of Deterrence Needs a Theory of Value

Imagine two intentional killers: a woman who, in anger, kills the sexual abuser of her young child; and a white man who, out of racist hatred, kills an African-American man for dating a white woman. How would we determine the appropriate punishments from an economic perspective? One thing that economics tells us is that to deter a crime of a particular sort the expected penalty — that is, the severity of the punishment discounted by the probability that it will be imposed — must exceed the gain to the offender.[1] Assume that the angry mother (and others in her situation) and the racist man (and others in his) get roughly the same psychic returns from their killings, and that the probability of detection in their two cases is roughly the same (close to one). Does economics imply under those circumstances that the two killers should receive the same punishment?

Not necessarily. We must determine in addition what value society gets from deterring these types of killings.[2] Society might attach relatively little value to the lives of child molesters; private acts of retaliation against them might make members of society happy and help to deter further child molestation. If so, then the social benefit of deterring deadly violence against child molesters

---

[1] *See* Gary Becker, *Crime and Punishment: An Economic Approach*, 76 J. of Pol. Econ. 169, 183-84 (1968); Jeremy Bentham, An Introduction to the Principles of Morals and Legislation 179, 184 (Hafner Library of Classics 1948).

[2] *See generally* Bentham, supra, at 171 (punishment is unjustified "[w]here it is *unprofitable*, or too *expensive*: where the mischief it would prevent would be greater than what it prevented").

might be low or even negative. The social benefit from deterring deadly attacks on African Americans, in contrast, might be high: members of society might attach a relatively high value to the innocent victims of such killings, and experience intense anger and fear in response to the racist motivations that underlie them. If so, then it might make sense economically for society to devote *less* (indeed, perhaps none) of its (inherently limited) punishment resources to punishing the angry mother, for the social benefit of deterring her type of killing is less than the benefit of deterring the racist man's.

Of course, matters could be otherwise. Maybe members of society are even less disturbed by the racist killings than they are by vengeance killings of child molesters, in which case economics might justify punishing the angry mother more severely. Or maybe they disvalue these killings equally, in which case the punishments should in fact be equal. The point, however, is that from an economic view the right level of punishment in these particular cases depends critically on how bad members of society take the underlying crimes to be.

And the same can be said in every other type of case. Unless we know whether and how much we disvalue a particular species of conduct, we can't determine whether the cost of deterring any particular amount of it is worth paying. Nor can we make a rational judgment about how to allocate punishment resources between crimes that impose different levels of harm.[3] The economic approach to deterrence, in short, presupposes a consequentialist theory of value.

2. The Theory of Value Is Essentially Political

That theory of value, moreover, is necessarily *political*. I mean this in two senses.

First, the theory of value is political in derivation. It's possible of course to derive such a theory from philosophy or theology or some other source that is putatively independent of politics. What's distinctive of economics, however, is its commitment to taking in-

---

[3] *See generally id.* at 181 ("*The greater the mischief of the offence, the greater is the expense, which it may be worth while to be at, in the way of punishment.*").

dividual valuations as given and aggregating them for purposes of determining what's socially valuable.[4] The market is thought to be the best mechanism for aggregating preferences when voluntary exchange is possible; democratically produced laws are when such exchanges aren't feasible or fail to make the actors internalize all the relevant costs. This isn't to say that democratic politics is a perfect mechanism for producing efficient laws, any more than it is to say that the market is the perfect mechanism for measuring efficient exchanges. But it is to say, for reasons that economists are well aware of,[5] that it's unlikely we can come up with institutional mechanisms for maximizing social welfare that do any better.

Second, the theory of value is political in nature. That is to say, the valuations that an economic theory of deterrence presupposes are open to perpetual contestation across all criminal-law-making institutions.

This is the case in part because of the imperfection of each lawmaking institution relative to the others. We might expect the judgments that a legislature would make about the respective harms associated with revenge killings of child molesters and racist killings of African-Americans to be an accurate reflection of the valuations of society's members generally; but a host of political dynamics, from inertia to interest-group rent-seeking, prevents us from concluding this with confidence. Individual juries are relatively immune from inertia and rent-seeking, but are less representative of the community as a whole and less sensitive to intensity of preferences than are legislatures. Judges, even when elected, are likely to be less responsive to the electorate than are legislators and less representative of the community than are juries; yet this very insularity from politics makes judges suitable representatives of

---

[4]*See* George J. Stigler & Gary S. Becker, *De Gustibus Non Est Disputandum*, 67 Am. Econ. Rev. 76 (1977).

[5]On the tendency of legislation toward efficiency, see Donald Wittman, The Myth of Democratic Failure: Why Political Institutions Are Efficient (1995); Gary Becker, *A Theory of Competition Among Pressure Groups for Political Influence*, 98 Q.J. Econ. 371 (1983); Sam Peltzman, *Toward a More General Theory of Regulation*, 19 J.L. & Econ. 211 (1976); on the tendency of judicial decisionmaking toward efficiency, see Richard A. Posner, The Economic Analysis of Law 23-24 (4th ed. 1994); George L. Priest, *The Common Law Process and the Selection of Efficient Rules*, 6 J. Legal Stud. 65 (1977).

citizens' "second-order" preference for the enforcement of higher principles (e.g., mercy and racial equality) that at least sometimes trump the immediate desires of the electorate.[6] Accordingly, the decision of no one of these institutions can be held forth as embodying the community's "true" aggregate valuation, which is in fact more likely to emerge from overlapping jurisdiction and competition among these institutions.[7]

In addition, the theory of value used to ground deterrence theory needn't remain static. Even if we could be confident that democratic institutions had for today accurately gauged the community's aggregate valuation of a particular crime today, it's possible those institutions would produce (or could be made to produce) a new result tomorrow as individuals' valuations changed. For that reason, too, the theory of value reflected in the decisions of democratic political institutions remains subject to perpetual contestation.

Another way to make the point is to say that nothing in the economic approach constrains anyone to accept the extant theory of value reflected in law, or prohibits anyone from advocating any particular new one. Race-of-the-victim disparities in juries' capital sentencing verdicts might reveal (as some have argued) that society places less value on the lives of African-Americans or poor people.[8] But that doesn't show (contrary to what others have suggested[9]) that welfare maximization justifies accepting such disparities. For if the people who serve on juries could be persuaded to alter the sentences they impose, or if legislatures or even courts could be persuaded to adopt rules of law that force juries to change their verdicts to eliminate the disparities, we could just as well say that the

---

[6] *See* Dan M. Kahan, *Democracy Schmemocracy*, 20 Cardozo L. Rev. 795 (1999).

[7] *See generally* Neil K. Komesar, Imperfect Alternatives: Choosing Institutions in Law, Economics, and Public Policy (1994).

[8] *See, e.g.*, Randall Kennedy, *McCleskey v. Kemp: Race, Capital Punishment, and the Supreme Court*, 101 Harv. L. Rev. 1388, 1441 (1988) (concluding that race-of-victim disparities in frequency of death sentence show that "in Georgia's marketplace of emotion the lives of blacks simply count for less than the lives of whites").

[9] *See* John R. Lott Jr., *Should the Wealthy be Able to "Buy Justice"?*, 95 J. Pol. Econ. 1307, 1311 n.7 (1987).

new sentencing outcomes reveal another, higher valuation of the relevant victims' lives, and that enforcement of *that* valuation is welfare-maximizing. Because theories of value are political in both derivation and in nature, anyone can argue for whatever theory of value they choose consistent with the economic theory of deterrence.

3. The Theory of Value Dominates Deterrence

My first two claims are sufficient to show that the economic conception of deterrence can never be fully autonomous of moral intuitions. To make optimal deterrence judgments, society needs a theory of value. That theory of value necessarily *is* the aggregation of citizens' intuitions about the reprehensibility of different crimes. Moreover, since society's theory is politically determined, nothing internal to the economic view precludes an individual from advocating that deterrence judgments be based on a theory of value derived from *her* particular intuitions.

I haven't yet demonstrated, however, that the economic conception of deterrence dissolves into intrinsic value retributivism. Intrinsic value retributivism holds that punishments should be based *solely* on intuitions about the reprehensibility of offenders' crimes. Even if informed by a theory of value constructed from such intuitions, the economic conception of deterrence could in principle generate punishments different from those society would arrive at if it were to apply common intuitions directly. Imagine that racist killers face a substantially higher likelihood of being caught and derive substantially less psychic utility from killing than do mothers who kill child molesters. Under those circumstances, the economic conception of deterrence might justify punishing the vengeful mother more severely than the racist killer even if society attaches *less* value to deterring the killing of child molesters than it does to deterring the killing of innocent persons who happen to be African-American.

What I want to show now is that deterrence theory lacks sufficient strength, conceptually or practically, to alter the path of common intuitions in this way. My final claim is that the external theory of value on which optimal deterrence judgments depend will always dominate the decisionmaking criteria internal to deterrence theory. This will be so, first, because individuals will always

have better information about their theory of value than they will about the factual issues distinctive of deterrence judgments; and second, because the disutility associated with overriding their moral intuitions will always cancel out any marginal social-welfare gains associated with crediting counterintuitive deterrence judgments.

a. Deterrence decisions under uncertainty. The economic conception of deterrence requires two sorts of determinations. One is how penalties of various sizes affect the incidence of an offense; this "penalty responsiveness" determination hinges on (among other things) the gain various classes of offenders derive from a particular kind of crime, and the likelihood that persons who commit such a crime will be apprehended and convicted.[10] The second determination is how bad various crimes are. As I've argued, this theory-of-value determination is necessary in order to assess whether the deterrence gains associated with a particular penalty are worth the cost of imposing it, and how limited punishment resources should be allocated across crimes.

A decisionmaker — whether a citizen, a legislator, a juror, or a judge — will always have better information on the valuation of a crime than on the components of penalty responsiveness. The reprehensibility of various types of crimes turns entirely on one's intuitive appraisals of them; to figure out, for example, which is worse — revenge killings of child molesters or racist killings of African-Americans — all one has to do is introspect. Introspection won't yield reliable answers, however, to the questions on which penalty responsiveness depends. Who really gains more (and exactly how much more) from his or her crime — the racist killer or the vengeful mother? Who is in fact more likely (and exactly how much more) to be caught and convicted? These are complex empirical questions, and answering them requires us to come to terms with possibly immense evidentiary and methodological difficulties.[11]

Because decisionmakers are bound to be more confident of their valuations of crime than they are of their assessments of pen-

---

[10] *See supra* note 1.

[11] *See generally* Franklin M. Fisher, *Multiple Regression in Legal Proceedings*, 80 Colum. L. Rev. 39 (1980).

alty responsiveness, their deterrence judgments will invariably be dominated by the former. This conclusion follows from our best psychological account of how persons do as a matter of fact make decisions under conditions of uncertainty, and is perfectly compatible with how individuals *should* as a normative matter make such decisions according to conventional economic theory.

Cognitive psychology tells us that individuals tend to conform their factual beliefs to their values. Individuals who place a high value on the autonomy of markets and other private orderings, for example, discount evidence that unregulated commerce poses a danger to the environment.[12] Those who think smoking is glamorous are less likely to believe empirical evidence that it causes cancer.[13] Individuals who see conformity to traditional, hierarchical social roles as entitling persons to status and respect are more likely to credit the claim that social deviancy promotes social disorder.[14]

Various mechanisms explain this phenomenon. Because values are likely to be more firmly held than factual beliefs, resolving uncertainty about the latter consistently with the former helps to avoid cognitive dissonance.[15] Social and cultural affinities reinforce this bias. Confronted with competing and offsetting empirical assertions that they aren't in a position to verify first hand, individuals rely on those whom they trust for guidance. The people they trust, unsurprisingly, are the ones who hold the same values they do, and who likewise instinctively resist the tragic conclusion that what's immoral might be beneficial, or what's moral, harmful.[16]

Studies of attitudes toward capital punishment confirm that these same dynamics influence beliefs about deterrence. Although the issue has been extensively researched, there isn't conclusive evidence one way or the other on whether the death penalty deters relative to life imprisonment. Nevertheless, citizens generally hold

---

[12] *See* Aaron Wildavsky & Karl Dake, *Theories of Risk Perception: Who Fears What and Why?*, 114 Daedalus 41 (1990).

[13] *See* Elliot Aronson, The Social Animal 185-86 (7th ed. ed. 1994).

[14] *See* Wildavsky & Dake, *supra* note 11.

[15] *See generally* Leon Festinger, A Theory of Cognitive Dissonance (1957).

[16] *See* Mary Douglas, Risk Acceptability According to the Social Sciences (1985); Mary Douglas & Aaron Wildavsky, Risk and Culture (1982).

confident opinions on this question, which correlate with whether they believe that the death penalty is a morally appropriate penalty for murder independently of its deterrent effect.[17] Indeed, experimental data show that individuals confronted with conflicting empirical evidence on the deterrent effects of the death penalty credit the evidence that confirms their prior beliefs and discount the evidence that conflicts with them.[18]

Values also have been shown to influence beliefs about the effect of gun control. Control supporters cite evidence that gun control reduces violence by disarming criminals, opponents that it *increases* it by disarming law-abiding citizens who could use weapons for self-defense. Which side's data one finds more credible depends on whether the prospect of widespread private gun ownership denigrates or affirms one's vision of a well-ordered society.[19]

There's no reason to expect decisionmakers (whether voters, jurors, legislators, or judges) to behave differently in other settings that require them to make deterrence judgments. If they are deciding how to allocate punishment resources between vengeful mothers and racist killers, they will predictably conform their speculations about the gain that these respective types offenders receive from offenses and the likelihood of their evading apprehension to their (more certain) beliefs about the intrinsic reprehensibility of these offenders' respective crimes.

Nor can they be faulted for doing so on the best economic account of how rational individuals *should* make decisions under conditions of uncertainty. Bayesian theory tells individuals how to update their probabilistic judgments (relating, e.g., to the various components of penalty responsiveness) as they encounter new information. Expected utility theory says that they should then use

---

[17] *See generally* Tom R. Tyler & Renee Weber, *Support for the Death Penalty: Instrumental Response to Crime, or Symbolic Attitude?*, 17 L. & Soc. Rev. 21 (1982); Phoebe C. Ellsworth & Samuel R. Gross, *Hardening of the Attitudes: Americans' Views on the Death Penalty*, 50 J. Soc. Issues 19 (1994).

[18] *See* Julian V. Roberts & Loretta Stalans, Public Opinion, Crime, and Criminal Justice 239, 242 (1997).

[19] *See* Dan M. Kahan & Donald Braman, *More Statistics, Less Persuasion: A Cultural Theory of Gun-Risk Perceptions,* U. Penn. L. Rev. (forthcoming 2003).

these judgments to determine the course of action (here the apportionment of punishment resources across different offenses) that yields the highest expected payoff.[20]

But while this strategy of decisionmaking tells individuals *what to do* with relevant and reliable information, it doesn't tell them *when* they should regard information as relevant and reliable. To avoid infinite regress, individuals must determine on some basis independent of rational decision theory what information merits inclusion in the rational-decisionmaking algorithm. Accordingly, a decisionmaker who is disposed to credit penalty responsiveness data when they cohere with her convictions about intrinsic reprehensibility and to discredit them when they don't isn't doing anything that a rational choice economist can object to. Indeed, if that's the form of information filtering that makes her "feel" best — because, say, it minimizes cognitive dissonance, or aligns her most comfortably with others with whom she identifies — then presumably she's doing exactly what she should be doing to maximize her utility.

b. The disutility of counterintuitive judgments. Imagine, though, that members of society devised a "value neutral" mechanism — perhaps an expert agency, manned by dispassionate and "worldviewless" econometricians — to gather information about the determination of penalty responsiveness internal to deterrence theory. Imagine further that the information suggested that it made sense on optimal deterrence grounds to punish some offense more severely than some other one that citizens would otherwise have deemed more intrinsically reprehensible. On the basis of the new information, the citizens might revise their intuitions about what level of punishment is just in the two cases. But if they did not, and if they remained convinced that the more intrinsically reprehensible offense should be punished more severely, would enforcing their intuitions be contrary to the normative prescriptions of economics?

No, because ignoring the citizens' collective preference for the intuitively just level of punishment in the two cases would reduce social welfare. If, on the whole, society decides that they'd rather

---

[20] *See* Howard Raiffa, Decision Analysis (1968).

punish the racist killer more than the vengeful mother notwithstanding the dictates of deterrence theory, then nothing in the economic analysis of law prevents them from accepting the trade off between optimal deterrence and the satisfaction of that intuition. Likewise if they choose to punish the mother more severely than the racist in the face of deterrence theory. Indeed, in both cases, their decision can be understood to reveal a new, upwardly revised assessment of the value of averting the types of killings committed by the offender whom they have chosen to punish more severely. For this reason, too, the external theory of value dominates the normative criteria internal to deterrence.

4. What All of This Amounts to

Put together, my claims demonstrate that the economic analysis of criminal law lacks the power to criticize intrinsic value retributivism. We can't make optimal deterrence judgments without a theory of value. Theories of value are always politically contestable; any argument about what's worse than something else can fairly be asserted and fought for. Because the decisionmaker will always have more information about her theory of value than about the other considerations relevant to optimal deterrence, her theory of value will always dominate her analysis. So although it's thereafter possible for her to reconstruct her conclusions in deterrence terms, those conclusions will in fact be derivative from the decisionmakers' intuitions about what crimes are more reprehensible than others. So we might as well just talk about what's more reprehensible than what, which is to say, we might as well talk intrinsic-value retributivism.