

POLISH ASSOCIATION OF LAW AND ECONOMICS

•
POLSKIE STOWARZYSZENIE EKONOMICZNEJ ANALIZY PRAWA

PSEAP WORKING PAPER
NO. 2
(DECEMBER 2007)

Jacek Czabański

*Why Consequentialists Should Be
Retributivists Too*

Dostępne na / available at: www.pseap.org

Polskie Stowarzyszenie Ekonomicznej Analizy Prawa (Law & Economics)
ul. Długa 44/50 pok. 407B
00-241 Warszawa
www.pseap.org
info@pseap.org

Jacek Czabański*

Why Consequentialists Should Be Retributivists Too

Abstract: The author argues that it is fully rational for consequentialist to be retributivists too. When there is no public law enforcement, rational victims will not be likely to punish the offenders, because the costs of punishment are higher than the potential deterrent gains for an individual. Therefore, rational offenders will not treat the threat of punishment as a credible one, and will continue to offend. Rational solution for potential victims is either to claim that it is retributive principle that they will follow, or to pay for the process of punishment in advance in order to avoid the temptation to save some money on the cost of general deterrence. In either way, it is a retribution principle that will be followed, although for consequentialist reasons. The author claims that both consequentialism and retributivism have their place within the theory of punishment, and one cannot be considered to be better than the other.

Streszczenie: Autor argumentuje, że również osoby preferujące celowościowe teorie kary powinny być retrybtywistami. Kiedy nie istnieje system publicznej egzekucji prawa, racjonalne ofiary nie będą skłonne do karania przestępców, ponieważ koszty wymierzania kar są wyższe niż potencjalne zyski dla jednostki w postaci zwiększonego efektu odstraszania. Tym samym, racjonalni przestępcy nie będą uznawali groźby kary za wiarygodną i będą kontynuować działalność przestępczą. Racjonalnym rozwiązaniem dla potencjalnych ofiar jest albo kierowanie się zasadą retrybucji, albo opłacenie procedury karania z góry, tak aby uniknąć pokusy zaoszczędzenia później pewnej sumy kosztem ogólnego efektu odstraszania. W obu przypadkach, stosowana będzie zasada retrybucji, aczkolwiek z celowościowych względów. Autor uważa, że zarówno teorie celowościowe jak i retrybtywne mają swoje miejsce w teorii kary i żadna nie może być postrzegana jako lepsza od drugiej.

Keywords: consequentialism, retributivism, theory of punishment

JEL: K14

* Ph.D. candidate, Centre for Law, Economics, and Institutions of the University of Turin, Italy. He may be contacted at jacek@czabanski.net. I thank Dr. Krystyna Pawłusiewicz for her valuable comments on the earlier draft.

I. Retributivism and consequentialism

There are two opposite camps within the theory of punishment. One is that of retributivists while the other is that of consequentialists.

The theory of retribution, which is now commonly assumed to be a dominant one in a public discourse of philosophy of punishment,¹ states, generally speaking, that criminals deserve punishment for wrongs they have done, and this reason is a necessary and sufficient condition for the infliction of punishment. Any other consequences, although may be welcomed if positive, are only additional.² A rival theory, consequentialism, claims that “it is right to punish criminals because doing so minimize the net level of suffering.”³

Because consequentialism looks for consequences, it is often described as forward looking, as opposed to retributivism which is backward looking.⁴ This was stressed by Bentham who argued that “general prevention ought to be the chief end of punishment as it is its real justification. If we could consider an offence that has been committed as an isolated fact, the like of which would never recur, punishment would be useless. It would be only adding one evil to another.”⁵

Such a view requires caution, though. Let us imagine a society in which there is only one period of time in which crimes may be committed and everyone is able to commit a crime only during that particular time, the *ex ante* threat of punishment is justified if it deters at least some of potential offenders. When *ex post* it happens that some people nevertheless have committed a crime, it is still worthy to punish them, even if, *ex hypothesis*, they cannot commit more crimes any more. This is not to deter them or others (which is impossible, as they will not have another opportunity), but it is to maintain the credibility of the initial threat made by law. This credibility is not necessarily tied with criminal law. In this case, for

¹ Dolinko (1997), p. 507.

² Sometimes, however, it seems that desert has to be supplemented by other conditions. For example, “punishment [ought] not [to] violate any non-forfeited rights of an offender” (Moore 1997, p. 173) or that punishment scale should be as low as possible, having in mind a necessity of the social order (von Hirsch and Ashworth 2005, p. 142). To this extend, desert is not a “sufficient” condition any more, although Moore argues otherwise (ibid.) Honderich (2006, p. 74) argues that retributivism is justified by providing grievance satisfaction. In that sense, it is consequentialist as well.

³ See Dolinko (1997), p. 507. Dolinko uses this description for a deterrence theory, but it applies to other consequentialist theories as well.

⁴ Simplifying, the opposition between retributivism and consequentialism is that in the former view we punish *because* a wrong has been done, while in the latter we punish *for* other wrongs not to be done. For such a distinction, see for example Rawls (1955), p. 5; Henderich (2006), p. 6.

⁵ Bentham (1843), p. 383.

example, there is no room for criminal law anymore (because by hypothesis people are unable to commit crimes any more), but the general notion of law enforcement survives. If promised punishment will not be carried on, any other rule of law may be perceived by people as unenforceable, and they will not feel obliged to obey any legal rules at all. Therefore, the forward looking aspect of consequentialism should be understood from *ex ante* perspective, i.e. before any crime has been committed, rather than from *ex post* perspective, when some crimes have already been committed and we only look for further consequences.

Arguably, in Kantian's situation of the civil society dissolution, it seems there is no need to punish anyone, which is of course against his retributive position. But this is not very likely that such a situation will occur. Moreover, if Kantian's postulate is to be interpreted that not only the last murderer has to be executed before dissolution (which is quite costless), but also that all crime mysteries have to be resolved before people can go on (which is quite impossible), this imperative considerably loses its intuitive power.⁶

In the terms of criminal policy, it may be argued that many crimes are committed by the first time offenders who are unlikely to continue their criminal activity, quite regardless a kind of punishment that will be inflicted on them.⁷ However, to say that because those people are unlikely to commit more crimes, no punishment is justified on the consequentialists ground would be premature. The *ex ante* perspective tells us that it might be the case that some people have been effectively deterred and did not commit crime because of the initial threat of punishment. Therefore, punishment should be inflicted on those who have committed crime to maintain the credibility of the system, even if those particular people are not likely to reoffend.

This story tells us that it is not always beneficial to calculate cost and benefits at every move, but it is better to set some set of rules to be followed later on. In such a way, a rational temptation for benefits in a short run will be defeated by a more beneficial long run perspective.

⁶ See Avio (1993), p. 263 ff, for that argument.

⁷ While many crimes are committed by first time offenders, a small fraction of offenders is responsible for gross number of crimes.

II. Rule utilitarianism

Rule utilitarianism refers to the long run perspective and forms basis for modern consequentialist theories of punishment, notably those of Hart's and Rawls'.⁸ In their views, the institutions of punishment have emerged because following such rules have good consequences. Goodness of the system makes it needless to consider every case on utilitarian grounds, or even makes it dangerous for the reason given above.⁹ Rawls stressed that the utilitarian rules of punishment, meant as a system, prevents from the temptation to do something beneficial in the short run only, for example punishing a person the judge knows to be innocent in order to maintain law and order in the community.¹⁰

More troubling question regards the question whether the utilitarian rules as such may allow punishing an innocent. For example, whether it is allowable on utilitarian grounds to punish someone that we only suspect is guilty. Therefore, it is a question whether institution of punishment may be used on people out of whom some may be at the end innocent. The general answer is, of course, yes. There is no possibility to punish guilty people only, because we do not have full information about things they have actually done. To say the system of punishment ought to be constructed in a way that precludes punishing an innocent, is to say that we cannot have any system at all.¹¹ Rawls tends to focus on the state punishment, but the conflict between the long run (rule utilitarianism) and short run (act utilitarianism) is as much acute in case of individuals as in the example above. Therefore, the claim that Rawls' distinction is valid only in the context of state punishment is misconceived.¹²

The distinction between the rules of practice and the practice under the rules allows Rawls to say that "the judge and the legislator stand in different positions and look in different directions: one to the past, the other to the future."¹³

⁸ See Hart (1968) and Rawls (1955).

⁹ The more common criticism of case by case considerations is that it leads to incoherent judgments. See Beccaria, "It is for this reason that we see the same court punish the same crime differently at different times, because it consults not the constant and fixed voice of the law, but the erring instability of interpretations." Beccaria (1995 [1764]), p. 15.

¹⁰ The alleged tendency to punish innocents, or to sacrifice an individual in sake of others, is of course one of the main argument against consequentialism and utilitarianism in general.

¹¹ The question is it morally permitted to have a system that inevitably leads to some positive number of punished innocents has not been particularly recognized within retributivism. But it still is a very troublesome question. See for example Smilansky (1990) p. 259, who argues that "...punishing' the innocent is in itself wrong, doing a great deal to avoid it is a matter of great importance – which must not be overridden *almost irrespective* of consequences" (emphasis added), and therefore any utilitarian attempts to relax standards of proof are morally unacceptable. But see Nozick (1974), p. 68, who correctly argues that there is no possibility to have a mistake-proof system of punishment.

¹² For that claim, see Zaibert (2006), p. 12-16.

¹³ Rawls (1955), p. 6.

People are generally sympathetic to retributive notion of justice. Recently, some game theory experiments reveal that people are ready to punish others for their non-cooperating behavior, even if it does not bring any benefits to themselves (because they will not meet that persons again), and even if the process of punishment is costly.¹⁴ Therefore, people are willing to pay something just for justice to be done, and apparently their own interests are not of the primary importance. However, in the long run, people in those social settings that allow imposition of punishment on non-cooperative members may be better off, so people will prefer (and migrate to) a punitive to non-punitive rules of game.¹⁵

III. Costly punishment and why we need retributivism

To see why consequentialists should adopt retributive principles consider a following example.

Let us imagine that there is no system of public law enforcement. Victims have to enforce the law on their own. When one has been robbed of, he has two choices. The first is to accept the loss and do nothing, the second is to chase the offender and try to recover the stolen goods. The chase is of course costly, and will be conducted only if expected benefits at least equal costs.

Suppose now that the chase has been successful and the thief has been caught. The stolen goods are recovered, and the question of punishment is to be decided. Inflicting punishment is costly for the victim, and she will rationally inflict the punishment only so far as the potential gains flowing from the deterrent effect outweigh the costs of inflicting the punishment. Therefore, we can expect that private punishment be more common in personal conflicts, when the offender's aggression has been directed toward a specific victim, and the victim can gain a lot in the sense of avoided future aggression.

However, when the offender chooses his victims quite randomly (for example, a street robber, or a burglar), the potential benefits of private punishment are quite small. Even if this particular offender will avoid this particular victim in the future, other offenders will not.

Therefore, the rational victim will refrain from the punishment, and the rational offender, having known that, will offend. Knowing that, victims can claim in advance that they will punish the likely offenders heavily, in order to deter them. This threat, however,

¹⁴ See for example Fehr and Gächter (2000) and Kahnemann et al. (1986) for evidence of retributive principle significance.

¹⁵ For evidence see Gürer et al. (2006).

lacks credibility, because the offenders knows that when the time comes, the victims will rationally decide not to punish them.

However, when people have intrinsic retributive beliefs that it is just to punish the offenders, regardless the possible outcomes (and particularly regardless the costs of punishment for themselves), the threat of punishment becomes reliable. Therefore, in retributive societies we can expect to have less criminals, and such systems are more efficient from the utilitarian standpoint. Therefore, it would be better for consequentialists to be retributivists, at least in the eyes of others.

IV. Fully rational consequentialists are retributivists too

Let us now suppose that people are fully rational, so we cannot convince others that we will punish an offender just because we share the retributive belief that evil deserves to be punished. How rational people can convince other rational people that they will not change their minds when the time to decide comes?

The answer has been given a long time ago by Homer. In the Iliad, Odysseus wants to hear the songs of Sirens, but he knows that when he has heard it, he will try to sail to them and will crash on the rocks of their island. Facing the situation when he in advance knows that he will choose the worse option, he decides to deprive himself of the freedom of choice, and orders his crew to tie him to the mast.

How then rational people can convince would-be offenders that they will be punished, even if it will be costly? Simply by paying for the process of punishment in advance. If there is no question of the cost of punishment in the future, it will be fully rational to punish the offenders, because not punishing them will not save any money. To avoid the temptation to save some money *ex post*, rational people will pay money to the fund which will cover the costs of punishing the offenders. While the reasons to establish this fund, and the institution of punishment in general, are completely utilitarian, those who are paid for inflicting punishment act on purely retributive grounds, i.e. they punish because wrong has been done. In fact, it would be contrary to the very reasons of establishing such an institution to allow it to consider any possible outcomes of its actions, and to modify its actions accordingly. There is no doubt that the present system of public law enforcement can be perceived as precisely such paid-in-advance system created by potential victims to overcome the problem of credibility of the private threat of punishment.

This is precisely what Rawls had in mind when he wrote that “the judge and the legislator stand in different positions and look in different directions: one to the past, the other to the future.”¹⁶

This is also why consequentialists should be retributivists. There is no possibility to be an efficient consequentialist (and to be an inefficient consequentialist is an absurd) but to apply the retributive principle in the process of punishment. However, while consequentialists turn to be retributivists at some stage, the nature of retributivism itself turns to be utilitarian. For the judge, or any law enforcement official, retributive justification should be sufficient. In fact, they are obliged to follow the principle of retribution by the very utilitarian reasoning that created the system of public law enforcement. However, legislators should not feel bounded by the principle of retribution, but rather should create the system according to the consequences it is likely to bring.

It explains why both retributivism and consequentialism have been so popular in the theory of punishment, and I hope it also tells us why one is not able to overcome the other in our thinking of punishment.

V. Bibliography

- Avio, K.L. (1993), “Economic, Retributive and Contractarian Conceptions of Punishment”, 12 *Law and Philosophy* 249.
- Beccaria, Cesare (1995 [1764]), *On crimes and punishments, and other writings* [ed. by R. Bellamy and tr. by R. Davies et al.], Cambridge University Press, New York.
- Bentham, Jeremy (1843), *The works of Jeremy Bentham, published under the superintendence of his executor, John Bowring*. Vol. 1. Principles of Penal Law. Edinburgh.
- Dolinko, David (1997), “Retributivism, Consequentialism, and the Intrinsic Goodness of Punishment”, 16 *Law and Philosophy* 507.
- Gürer, Özgür; Irlenbusch, Bernd; Rockenbach, Bettina (2006), “The Competitive Advantage of Sanctioning Institutions”, 312 *Science* 108-111
- Hart, H.L.A. (1968), *Punishment and responsibility: essays in the philosophy of law*, Oxford University Press.

¹⁶ See note 13 supra.

- Honderich, Ted (2006), *Punishment. The Supposed Justifications Revisited*. Pluto Press.
- Kahneman, Daniel, Knetsch, Jack L., Thaler, Richard H. (1986), "Fairness and the Assumptions of Economics," 59 *Journal of Business* 285.
- Moore, Michael (1997), *Placing Blame. A General Theory of the Criminal Law*, Clarendon Press, Oxford.
- Nozick Robert (1974), *Anarchy, State, and Utopia*. Basic Books.
- Rawls, John (1955), "Two Concepts of Rules," 64 *Philosophical Review* 3-32.
- Seidman, Louis Michael (1984), "Soldiers, Martyrs, and Criminals: Utilitarian Theory and the Problem of Crime Control," 94 *Yale Law Journal* 315 – 349.
- Smilansky, Saul (1990), "Utilitarianism and the 'Punishment' of the Innocent: The General Problem," 50 *Analysis* 256-261.
- Von Hirsch, Andrew and Ashworth, Andrew (2005), *Proportionate Sentencing. Exploring the Principles*. Oxford University Press.
- Zaibert, Leo (2006), *Punishment and Retribution*, Ashgate.