On the Evaluative Criteria for Laws: A Synoptic Overview and Some Possible Avenues for Research*

Satish K. Jain
Email: satish.k.jain@gmail.com

Before I begin, I would like to express my thanks to NALSAR for organizing this Conference and for their kind invitation to make this presentation. In this talk I will briefly discuss: (i) Some non-efficiency criteria that can be analysed within the framework of law and economics; (ii) Some criteria for laws that cannot be analyzed within the framework of law and economics; and (iii) Some possible avenues for research regarding these criteria.

The method that economists have developed for analysing rules in general, and legal rules or laws in particular, consists of asking, for every specification of relevant conditions, what actions purposive individuals would undertake within the framework of the rule in question; and of determining what outcome would result as a consequence of the totality of actions undertaken by the individuals. Once such an exercise has been done for every possible specification of relevant conditions, one knows the set of all outcomes which are possible through the interaction of rational individuals within the framework of the rule in question. In the economic analysis of rules and institutions the most important question that one asks is whether the rule or institution has the property of invariably yielding efficient outcomes. Thus when one uses the expression 'economic analysis' in the context of study of institutions or rules, one usually has in mind both the method mentioned above, which can be called the economic method as a shorthand expression, as well as determination of efficiency or otherwise of outcomes generated under different specifications of the relevant conditions. It is important to note that the two elements of economic analysis are logically distinct and separable. There is no reason why the set of possible outcomes cannot be analysed from the perspective of criteria or values other than efficiency. In what follows I discuss three important non-efficiency criteria, namely, restitution, retribution, and deterrence.

*Presented at the 'Law and Economics: Theoretical and Empirical Explorations' Conference held at NALSAR University of Law, Hyderabad, 19-20 January 2019.

The criterion of restitutive justice requires that victim be compensated for loss in a manner such that after the compensation he/she is restored to pre-loss position or an equivalent position. If a law or legal rule is such that it invariably makes the victim indifferent between pre-loss position and post-loss position with compensation, the law or the rule can be said to possess the attribute of being exactly or fully restitutive in character. Similarly one can define a rule to be over-restitutive if under it the victim is invariably made better-off in the post-loss position with compensation compared to the pre-loss position; and under-restitutive if the victim is invariably rendered worse-off in the post-loss position with compensation compared to the pre-loss position. Needless to say, these three categories are not exhaustive; and there are many rules which do not belong to any of the three categories. For instance, the rules which make restitution contingent on factors other than the amount of loss suffered by the victim would in general belong to none of the three categories. While from the perspective of justice under-restitution may not be acceptable, over-restitution may not be found objectionable. Over-restitution, however, may be problematic from a different perspective. If a law systematically over-restores it is possible that it may result in victims developing perverse preferences for wrong-doing by perpetrators. Ideally, legal rules should aim for exact restitution.

While the restitutive justice focuses on the victim, the retributive justice relates to the perpetrator. If under a rule punishment to the perpetrator is exactly equal to harm to the victim the rule would be called an exactly retributive rule; if punishment is greater than harm then over-retributive; and if punishment is less than harm then under-retributive. Unlike the notion of restitution, the idea of retribution is not completely unambiguous. The ambiguity arises because it is not entirely clear whether for the purpose of calculating the punishment one should compare the pre-crime position with the post-sentence position or compare the post-crime but pre-judgment position with the post-sentence position. Suppose individual A takes away from individual B something valued at 100. If the thing is restored to individual B but no additional punishment is given to A then individual A's position in the post-sentence position becomes the same as his pre-crime position implying zero retribution under the first and stricter interpretation; but if we compare his post-crime but pre-sentence position with the post-sentence position then the retribution works out to be equal to the initial loss of the victim implying full retribution. In this particular instance, it is clear that our intuitive understanding of retributive justice would not regard punishment in this case as full retribution. This might suggest that the stricter definition perhaps captures our intuitive notion of retribution better. But consider a case where individual A has killed individual B; and individual A has been given capital punishment. Most people would regard it as a case of exact retribution. Here, it is clear that the comparison is between post-crime pre-sentence position and post-sentence position.

Our intuitive conception of retribution, unlike in the case of previous example, is in harmony with this comparison. Thus, it appears that both ways of defining retribution may be relevant in different contexts.

The idea of deterrence is also quite complicated. There are two different senses in which the concept of deterrence can be defined. One of these can be termed as absolute deterrence. If one assumes that potential perpetrators are rational, i.e., purposive then, if the quantum of punishment is greater than the benefit to the perpetrator from wrong-doing then the act involving the wrong-doing would not be undertaken at all. Thus absolute deterrence requires penalty to be greater than the benefit from wrong-doing. But there is a different sense in which one could talk of deterrence. If there is an activity which is beneficial to the perpetrator but harmful to the victim, one might take the position that it should not be undertaken if harm is greater than the benefit but should be permitted to be undertaken if the benefit is greater than the harm. Implicit in this way of defining deterrence is the idea that only those activities which in the aggregate are harmful should be deterred.

Whether one adopts the strict version or the weak version, under-deterrence seems to be undesirable. The question of the desirability or otherwise of over-deterrence crucially depends on which of the two conceptions of deterrence is adopted. If one wants to deter if and only if in the aggregate the activity is harmful, then clearly over-deterrence is undesirable. On the other-hand if one wants to deter absolutely then the extent to which one deters overly would be a matter of indifference.

As a first step for discussing these criteria systematically one can make use of the following simple framework.

Let A be a subset of $\mathcal{A} = \{(l, g) \mid l \geq 0, g \geq 0\}$; J a subset of $\mathcal{J} = \{(c, p, t) \mid c, p, t \text{ are numbers}, c \geq 0, p \geq 0\}$. A (legal) rule f is a function from A to J. A harmful act or crime would be denoted by an ordered pair of numbers (l, g), $l \geq 0, g \geq 0$; where l denotes harm to the victim(s) and g gain to the criminal(s) or the perpetrator(s) of the harmful act. Judgment or the court decision would be represented by an ordered triple of numbers (c, p, t), $c \geq 0, p \geq 0$; where c denotes compensation to the victim for the harm, p punishment to the criminal or the perpetrator of the harmful act, and t the net transfer to the rest of the society. The units of measurement of l, g, c, p, t would be assumed to be identical.

To ensure invariance with respect to unit of measurement, only those rules are to be

considered that satisfy the following property:

$f : A \mapsto J$ is such that $(\forall(l, g) \in A)(\forall(c, p, t) \in J)(\forall\lambda > 0)[f(l, g) = (c, p, t) \rightarrow f(\lambda l, \lambda g) = (\lambda c, \lambda p, \lambda t)]$.

We define a judgment to be (i) exactly restitutive iff $c = l$, overly-restitutive iff $c > l$, and under-restitutive iff $c < l$; (ii) deterring iff $p > g$, and under-deterring iff $p < g$; (iii) exactly retributive iff $p = g + l$, over-retributive iff $p > g + l$, and under-retributive iff $p < g + l$; (iv) exacting iff $p = l$, over-exacting iff $p > l$, and under-exacting iff $p < l$; (v) efficient iff $[(g - l > 0 \rightarrow g - p > 0) \wedge (g - l < 0 \rightarrow g - p < 0) \wedge c + t = p]$; and (vi) closed with respect to the harmful act (crime) iff $t = 0$; and open iff $t \neq 0$.

If every judgment under a rule f possesses a certain attribute; as a short hand expression we will speak of the rule having the attribute in question. For example, if we have $(\forall(l, g) \in A)[c \geq l]$ then $f : A \mapsto J$ will be called a restitutive rule.

It is fairly clear from the way the ideas of restitution, retribution and deterrence have been defined here that one cannot expect any rule to be uniformly satisfactory from the perspective of all the criteria. Consequently the choice of laws or rules in the main would depend on which criterion is considered of paramount importance in a particular context or to be more precise on the hierarchy of criteria which is considered appropriate in a given context.

The different branches of law differ widely with respect to the three criteria being considered here as well as with respect to other criteria. It is important to analyze the important branches of law in terms of these three criteria and also discuss whether the differences are in consonance with our intuitive feelings of the idea of justice.

The structure of criminal law is such that by and large one can say that the attribute of being retributive holds for it. While it is not difficult to find examples of both grossly under-retributive and grossly over-retributive laws, given the fact that the principle of proportionality, that is to say the principle that the punishment should be proportional to the gravity of crime, commands universal acceptance, one should expect a significant proportion of criminal laws to be by and large neither too over-retributive nor too under-retributive. In any case the objective of retribution clearly figures most prominently in the design of criminal laws. For crimes, where harm to the victim exceeds gain to the perpetrator, an overly or exactly retributive law would also be absolutely deterrent, as by committing the crime the perpetrator would lose more than what he would gain. But for

crimes, where harm to the victim is less than the gain to the perpetrator, the deterrence aspect depends whether the law is retributive in the strict sense or the weak sense. If the law is retributive in the strict sense then it is always absolutely deterrent regardless of the quantum of the harm as long as there is some harm. However, if the rule is retributive only in the weak sense then in case the harm to the victim is less than the gain to the perpetrator, law would only be deterrent in the weaker sense. Consequently if the society wants criminal law to be both at least fully retributive and absolutely deterrent then the punishment must be at least as great as the larger of the two, harm to the victim and gain to the perpetrator; and in case gain to the perpetrator is at least as large as loss to the victim then the punishment must be greater than the gain to the perpetrator. It should be noted that in general it is not possible to have rules to be both absolutely deterrent and exactly retributive if retribution is defined in the weaker sense.

A remarkable feature of the criminal law as it has evolved is that the restitutive justice is altogether missing as an attribute. The structure of criminal law is such that after the act of crime the victim becomes quite peripheral to the unfolding of the legal process. The society as a whole in the form of the state becomes the aggrieved party; and what it wants from the trial is retribution against the criminal and not restitution for the victim. While it is understandable that the position of the plaintiff in criminal cases should be occupied by the state as acts which constitute criminal offences normally are violations of those values which the society considers to be of greatest significance for its survival or for its raison d'être; and that retribution and deterrence should occupy the top places in the hierarchy of attributes in the design of laws dealing with such offences. But what is not clear is why the restitution should not figure at all. It is not as if the restitutive considerations were in inherent conflict with retributive and deterrence objectives.

From the perspective of restitution the structure of property law is the polar opposite of that of criminal law. While in criminal law the restitutive considerations are conspicuous by their absence, they constitute the predominant features of the property law. In general, perpetrators of harm have to fully compensate for loss to victims. Compensatory damages are the norm for harm to property. Also, interestingly the provision of injunction ensures that by and large there would be absolute deterrence rather than the weaker version of deterrence. As in the case of property, harm to the victim and the gain to the perpetrator tend to be more or less the same, restitution by the perpetrator constitutes retribution also. Thus we see that it is in the context of the property law that all the three criteria that we have been considering are to a large extent satisfied.

The structure of contract law has some interesting similarities as well as differences

with property law. In case of breach of contract if the promisee is awarded reliance damages, with the notion of reliance damages broadly conceived, then after the award of damages the promisee would be restored to the pre-contract position. Consequently, reliance damages constitute exact restitution. The penalty of reliance damages also results in deterrence and exact retribution, both in the weak sense. However, the norm in contract law is not reliance damages but expectation damages. Award of expectation damages makes the promisee as well as off as she/he would have been had there be no breach of contract. If the contract is performed the promisee would be at least as well-off as she/he was in the pre-contract position. Thus, expectation damages would result in exact restitution or over-restitution. The deterrence would be the same or greater than what it would be in the case of reliance damages; and a similar conclusion holds with respect to retribution. The major difference between contract law and property law pertains to the fact that, unlike property law, the contract law does not deter absolutely. If the gain to the party breaching the contract is greater than the harm to the other party because of breach then the remedy of expectation damages ensures that breach would take place. Thus contract law deters only in the weak sense.

The analysis of tort law is more complicated as it uses rules for apportioning liability based on two quite different ideas of strict liability and negligence. These two have very different implications with respect to the three criteria that we have been considering. Under strict liability the wrong-doer is fully liable for the harm caused to the victim. Thus under the strict liability there is exact restitution as well as exact retribution (in the weak sense). With respect to deterrence, the rule of strict liability is akin to reliance damages in ensuring only the weak deterrence. If the gain to the tortfeasor is greater than the harm to the victim, the tortfeasor would commit the tort and pay full compensation. Only when the harm is greater than the gain that the strict liability would deter altogether.

Under the rule of negligence, the tortfeasor is liable for harm if and only if the tortfeasor is at fault, that is to say, the level of precaution taken by the tortfeasor is less than the due care level. A rational tortfeasor would always take due care; and consequently would not have to pay for the harm to the victim. Thus, if everyone behaves rationally then, under the rule of negligence, there would be no restitution, no retribution and no deterrence.

In the law and economics literature quite general results have been obtained regarding the characteristics of rules that invariably give rise to efficient outcomes. It would be interesting to study the structure of restitutive, retributive, and deterrent rules along similar lines as has been done in the case of efficiency. In this connection it is important to remember that the efficiency analysis of rules has shown that the characterization of

efficient rules crucially depends on how the specification of variables is done. For instance, in a model where the victim cannot take any care for reducing the expected harm, the strict liability is efficient; but not otherwise. A similar pattern can be expected in the case of the three criteria under consideration as well. That is to say, in all likelihood, the characterizations of rules possessing these non-efficiency attributes would depend on how l, g, c, p are specified.

When one analyses legal rules by the economic method as outlined earlier, where they are treated as rules of the game, one is necessarily taking an approach whose focus is on consequences. Legal rules of course can be, and are, analysed from other perspectives. One important way in which legal rules or laws can be normatively analysed is to determine which values are reflected or embodied in the text of the rules. For instance, if a law specifies as punishment a monetary fine for one group of people and imprisonment for another group of people for the same crime, most people would have no hesitation in terming the law as unjust. In this connection it is important to note that the existence of a rule which specifies unequal treatment for different sets of people does not necessarily imply existence of legal cases involving unequal treatment. This can happen for any number of reasons. Suppose for instance, the punishment of imprisonment is so harsh compared to the likely benefits of the activity proscribed by the legal rule in question that it proves to be an absolute deterrent; and consequently the legal rule is violated by no person belonging to the category for which imprisonment is the punishment for transgressing the provisions of the rule. Thus all violators would be from the category of those for whom the punishment is levying of a fine. Thus as far as the actual outcome is concerned all violators of the rule would have been treated equally. In this example it is evident that the absence of cases with unequal treatment would not make the rule in question any less unjust. For the purpose of determining which values are reflected and embodied in legal rules the textual analysis of a body of laws would have to look at not only individual rules, subsets of rules and the entire set of rules but also their interrelationships. For instance, each of the two legal rules dealing with transgressions of similar gravity taken singly may be fine; but together may be objectionable on the ground of disproportionate punishments for crimes of more or less similar magnitude.

One of the most important attribute for any set of rules is that of internal consistency. If a set of rules suffers from logical inconsistency then the outcomes can become uncertain and arbitrary. The internal consistency for the body of laws is even more important as the absence of it can lead to miscarriages of justice. As internal inconsistency of a set of legal rules in general can be expected to lead to arbitrariness of the outcomes; one likely consequence of it would be excessive litigation. Although excessive litigation can be

there for any number of reasons, three of them, including the internal inconsistency, are particularly important from a theoretical perspective. As it is impossible to draft rules in a manner such that there would be no possible scenario where the application of the rule would be ambiguous, some litigation is possible no matter how well-crafted the rule might be. Having said this, it is also clear that a well-drafted rule would generally take care of most of the relevant and likely scenarios, and thus would be unlikely to lead to much litigation. Therefore one major source of excessive litigation lies in poorly-drafted enactments. Another source of excessive litigation lies in the divergence between the objectives for which laws are enacted and the objectives that are actually realized. For the proper functioning of any institution and the realization of the the values or objectives for which the institution supposedly exists it is crucial that the tasks that are required for the proper functioning of the institution are performed in the way they are supposed to be performed. However, in general, it is unlikely that interests of individual performing the institutional tasks would be perfectly aligned with what is required of them for the proper functioning of the institution. Thus the actual values or objectives that are realized through the instrumentality of the institution in question would depend on whether individuals by and large act in their self-interest or perform their assigned tasks properly. The desired objectives and values for which the institution exists are unlikely to be realized if most individuals act in their self-interest rather than act the way demanded by the institutional requirements. In the context of any institution, wide divergence between the desired objectives and realized objectives can lead to serious societal problems. The divergence in the context of legal institutions can be expected to lead to excessive litigation.

One necessary consequence of excessive litigation is clogging of courts and the consequent delays in the disposal of cases. Excessive delays in the disposal of cases have the potential of corroding the system of checks and balances. Thus, for any legal system, in case of clogging of courts, it is of utmost importance to find out the reasons for it. Among the three factors that have the potential of giving rise to excessive litigation discussed above, two of them solely relate to the texts of laws and their interrelationships and therefore can be studied theoretically. The study of behavioural divergences from the way the individuals are supposed to function of course would require empirical analysis.

It is clear that the normative analysis of laws along the lines indicated here can go a long way in enhancing our understanding of the legal system as well as indicate the direction in which reforms should to be undertaken.

Thank you !