Aristotle and Optimal Deterrence

Alessandro Romano1

1 LUISS Guido Carli, Rome, Italy

Correspondence: Alessandro Romano, LUISS Guido Carli, Rome, Italy. E-mail: aromano@live.it

Received: June 22, 2013   Accepted: July 31, 2013   Online Published: August 30, 2013
doi:10.5539/jpl.v6n3p146          URL: http://dx.doi.org/10.5539/jpl.v6n3p146

Abstract

The central role of corrective justice in shaping private law has vigorously been challenged by the law and economics movement, and its concept of optimal deterrence. The fierce rivalry between the two competing theories seems to be at a dead end since scholars on each side appear to be deaf to the arguments raised by their counterparts. However, building on Kantian notion of conceptually sequenced ideas, it is suggested that the two theories are not only perfectly compatible, but also necessary complements of each other. Law needs justice, justice needs deterrence, and deterrence without justice is an empty and incoherent exercise.

Keywords: corrective justice, deterrence, Kantian rights, efficiency

1. Introduction

“Prior to 1970 the debate on the purposes underlying the law of torts was either empty or banal” (Schwartz, 1996 p.1802). However, in the past decades the situation has drastically changed and a fierce battle has exploded between the champions of the Aristotelian “corrective justice” (Fletcher, 1972; Epstein, 1973; Coleman, 1982; Weinrib, 2012) and law and economics scholars (Calabresi, 1970; Posner, 1972; Kaplow & Shavell, 2009).

In spite of a few attempts of reconciliation (Chapman, 2001; Geistfeld, 2001 and 2009) it appears that there is an enormous gap between the opposing doctrines, as they seem to imply completely incompatible policies.

Law and economics scholars would chase optimal deterrence; it is conventional wisdom that it can be achieved only if damages are equal to the harm times the inverse of the probability that compensation is due (Polinsky & Shavell, 1998). The obvious implication is that the use of punitive damages should be widespread, since such probability is strictly smaller than one.

With regards to sanctions, one of the core ideas advocated by Becker is that if sanctions are monetary and individuals are risk neutral in wealth, then optimal sanctions are maximal (Becker, 1968; Polinsky & Shavell, 1999). At a first glance it might appear that, from an economical perspective, the obvious solution to achieve optimal deterrence at the least cost would be to increase the magnitude of the fines and to introduce (or enhance, depending on which side of the ocean we stand) that intriguing creature known as punitive damages.

Conversely, it is not possible to offer a single definition of corrective justice and hence for now the focus will be on the original definition offered by Aristotle. According to him, corrective justice involves the notion of balance, or equipoise, between two individuals (Aristotle trad., 1962). Torts can be considered as transactions that alter this balance; corrective justice aims at righting the scales. As corrective justice sees remedies as a mean to undo the wrongs (Weinrib, 2002), it is straightforward that the ideal solution would be to have certain compensation that is equal to the harm (Note 1).

Using a more formal language, a superficial look at deterrence theory might induce one to think that the probability of detection by public authorities should tend to zero (with the consequent introduction of enormous fines), and compensation should always exceed the harm. Neither of these two policies is implied by corrective justice.

Not only the opposing doctrines have completely diverging policy implications, they are usually assumed to be absolutely incompatible since they rest on opposite axioms.

However, starting from the Kantian argument of conceptually sequenced ideas, it will be suggested that, not only they are not mutually exclusive, but they are necessary complements. The one without the other cannot offer a satisfying description of tort law as a whole.
Two examples will be introduced to ease the exposition: the first involves a case of product liability, the second a hard core cartel.

2. Deterrence Theory

The idea of sanctions (or more generally punishment) to deter unwanted behavior has a millenarian history (Lebow, 2007), but an economic formalization is owed to the pioneering works from Becker (1968), Calabresi (1970) and Posner (1972).

According to the economic theory of deterrence, a potential injurer will be optimally deterred only if his expected liability is equal to the potential harm that his conduct might cause. In mathematical terms optimal deterrence will be achieved if:

\[
P_f M_f + P_d D = H
\]

Where \( M_f \) is the magnitude of the fine attached to the wrongful action, \( P_f \) is the probability that such a fine will actually be inflicted on the injurer, \( D \) is the amount that the injurer will have to pay to compensate the victim, \( P_d \) is the probability that compensation will be due, and \( H \) is the harm suffered by the victim.

A few key points should be noted. In the first place, this very general formulation is extremely flexible and can be adapted to a very diverse set of cases. A driver that is considering the opportunity of speeding will take into account both the expected fine and the expected liability from a possible accident, thus the equation (1) appropriately describes his sets of incentives. Clearly this is not always the case, as often no sanction is attached to a conduct that generates a tort. Under these circumstances the expected fine will be equal to zero and the total expected liability (Note 2) will simply be equal to the expected damages. Optimal deterrence will be achieved for \( P_d D = H \).

In the second place, the right hand of the equation represents the harm suffered by the victim, instead of the benefit gained by the injurer. In fact, if the only concern is to maximize overall efficiency, a certain conduct is desirable whenever the benefits are greater than the expected harm. If the expected liability of the injurer is exactly equal to the expected harm, he will engage in the activity if and only if its potential benefits are greater than that. The injurer will be able to compensate the victim, while still keeping some of the benefits derived from his conduct. In other words it is possible to achieve, at least potentially, a Pareto improvement.

In the third place, it can be observed that when \( M_f > 0 \) and \( P_f > 0 \), not the expected fine (\( P_f M_f \)), neither the expected compensation (\( P_d D \)) should be equal to the harm, but their sum should be.

A last point that will be noted is that this mathematical representation might be misleading in many contexts, including our antitrust example. In fact, at least at a theoretical level (Note 3) every conduct that is deemed to be anticompetitive generates a deadweight loss (DWL) (Note 4), on top of the harm inflicted to the victim. To achieve optimal deterrence the formula should be rewritten as follows:

\[
P_f M_f + P_d D = H + DWL
\]

If the injurer is not induced to consider also the dead weight loss, but only the direct harm inflicted to consumers, (Note 5) it will impose an externality on the society and engage in inefficient conducts.

The condition imposed by (2) is necessary but not sufficient to maximize social welfare, especially in case of bilateral accidents. In order not to give victims excessive (or insufficient) incentives to sue (2) should be divided into the following equations:

\[
P_f M_f = DWL
\]

\[
P_d D = H
\]

It is straightforward that if (2) is verified, (3) and (4) have to hold simultaneously. The reason to impose these additional conditions can be understood focusing on (4). In fact, if \( P_d D > H \) the victim would potentially be better off if the accident takes place, and hence, he will have no incentives to take any precaution (and might even actively try to increase the probability of an accident). On the other hand, if \( P_d D < H \) the victim is forced to internalize some of the expected losses caused by the injurer, thus will adopt an excessive level of care or a sub optimal activity level.

In other words, in order to induce both parties to behave optimally the expected fine should be equal to the deadweight loss (as defined here) caused by the conduct, while the expected compensation owed by the injurer should equal the expected damages.

As stated above, if the only goal to be achieved is economic efficiency, the seminal paper by Becker implies that the \( P_f \) should tend to zero, while the monetary fine should tend to infinity (Becker, 1968). At the same time, since
$P_d$ is strictly less than one, $D$ should always exceed $H$.

Such a solution might appear extreme, and in fact law and economics scholars have identified many reasons why Becker’s claim only works under certain conditions and up to a certain threshold. A survey of these considerations lies outside the scope of this work (see Friedman & Sjostrom, 1993); here it suffices to say that economists focus on the total expected liability faced by the injurer. Moreover, within certain limits, Becker’s claim is widely regarded as correct; therefore an economist would not hesitate to call for higher sanctions whenever the probability of detection of a certain conduct is low, either because of the nature of the conduct, or due to the budget constraint of the enforcer. At the same time, it is completely dominant in the economic literature the idea that compensation should equal the harm times the inverse of the probability that compensation is due (Polinsky & Shavell, 1998).

3. Corrective Justice

Just like deterrence theory, corrective justice has an illustrious tradition and its roots go as far as ancient Greece. Unlike deterrence theory, though, it is impossible to offer a description that would be accepted by all the scholars that consider it to be the cornerstone of private law. For this reason, only the aspects that are relevant to the discussion at hand will be underlined.

The first, and undoubtedly most important, characteristic of corrective justice is the cardinal connection between the two parties, and hence, its focus is invariably on binary relationships. Corrective justice inevitably deals with pairs of actors, generally labeled as the doer and the sufferer of an injustice. The position of the doer and sufferer are inextricably tied together, because “the doing and the suffering of an injustice are the active and passive correlates of each other” (Weinrib, 1997 p. 41). In other words the indissoluble link between the doer and the sufferer is the most characteristic trait of corrective justice. In the second place, as Weinrib writes, “the plaintiff’s [sufferer] suit is an attempt to vindicate a right that the defendant [doer] has unjustly infringed” and the “remedy rectifies the injustice and thereby reflects its structure and content” (Weinrib, 2000). A very important point is that rectification is considered to be aimed at undoing the wrong, therefore the remedy has to be a response to the factors that are constitutive of the injustice (Weinrib, 2000). To use Aristotle terminology, corrective justice aims at achieving fairness and equality (to ison) and can be expressed in terms of equality of quantities, explicitly defined in mathematical terms by the Philosopher. The equality should not be intended in a literal sense, but in a notional one. Equality consists in persons having their own and the holdings of the parties prior to the interaction constitute the relevant baseline. In other words, before the interaction the doer (A) and the sufferer (B) will have a certain endowment that is assumed to reflect equality:

$$A = B \quad (5)$$

After the interaction the doer will now have something that belongs to the sufferer, and hence:

$$A > B \quad (6)$$

As stated above, corrective justice aims at righting the scales by bringing A and B in the situation described by (5). Furthermore, the victim should ideally be in the same situation in which she was before the interaction. We write:

$$A_{id} = A_{oi} = B_{id} = B_{oi} \quad (7)$$

To denote that at the time $t_1$, the sufferer should not only be equal to the doer at time $t_1$, but should also be equal to herself at time $t_0$.

It is crucial to note that this equality cannot be considered as merely referring to wealth, yet money has to be an -at least imperfect- substitute of the relevant variable, whatever that is. It would make very little sense to claim that the harm should be rectified by a wealth transfer if this was not the case.

If money is not considered an (imperfect) substitute of the relevant variable, it becomes impossible to fit in the corrective justice framework the fact that in every developed legal system the largely predominant route taken by courts to compensate the sufferer is to attach a monetary value to the injustice. If we are to defend the idea that tort law is shaped by corrective justice, we should embrace the idea that wealth is the closest substitute of the relevant variable that can be manipulated by courts. It is possible to imagine a wide range of different remedies (e.g. the doer could be forced to work for the sufferer for a certain time or to provide a certain service), however if damages are the (quasi) universally accepted remedy we must conclude that money are in fact the best substitute available of the relevant variable.
4. An Impossible Marriage?

The two “schools” appear to be destined to a never ending fight. The scholars on each side seem to be much more concerned with refining their own theory, than with finding a common ground to debate. Not surprisingly the more the two theories are refined, the more they become abstract and impenetrable to the scholars of the other faction. Economists have developed models that are extremely complex and sophisticated, but that at times appear to have a very low degree of adherence to the real word. On the other hand, corrective justice theorists are becoming more and more interested in the fascinating, yet challenging work of Immanuel Kant. From this perspective economists are not alone in doubting that the framework developed by the German Philosopher can be useful to understand the modern world, as even many legal scholars appears to be skeptical. The idea that law can be presented as lean, minimal and self-contained seems to be in sharp contradiction with the pluralism advocated by many contemporary scholars (Weinrib, 1987).

The first attempt at lessening the gap between the two theories is the already cited paper from Gary Schwartz (1997) but even his work only created a small chink in the wall that separates deterrence theory and corrective justice.

The crucial point raised by Weinrib in reply to Schwarz is that corrective justice is intrinsically focused on the binary relationship between the doer and the sufferer, whereas this relationship hardly has any relevance for deterrence theorists. In the words of Posner it makes economic sense to take money from the defendant in order to induce him to take cost-justified precautions, “but that the damages are paid to the plaintiff is, from an economic standpoint, a detail” (Posner, 1998, p.78).

Furthermore, deterrence theory introduces exogenous goals like loss spreading that are completely incompatible with the framework developed by corrective justice theorists. “It seems that loss spreading is over-inclusive as a goal of tort law, and the idea that money should be exacted from some for the benefit of others in order to spread the burden of a catastrophic loss as lightly and as widely as possible is as pertinent to non-tortious injury as to a tortious one.” (Weinrib, 1995, p.37) If the inspiring idea is to exploit the concept of diminishing marginal utility in order to maximize social welfare, there is no reason to even introduce tort law, as social insurance and progressive taxation appear to be more appropriate to pursue this goal (Calabresi, 1970). From this perspective the efficacy of tort law is strongly limited by the requirement of causation that is indeed heterogeneous to the idea of loss spreading.

The burden of proving why tort law should be an appropriate mechanism to achieve loss spreading then lies on law and economics scholars. For instance, it could be argued that a centralized system of social insurance might be extremely cheaper in terms of administrative costs than the elephantine apparatus necessary to tort law. At the same time, incentives could be preserved calibrating insurance premiums on the level of risk of each individual. Given the focus on overall utility adopted by economists, it appears rather paradoxical that only plaintiffs who are lucky enough to be harmed by a wealthy defendant have the privilege of not having to bear a concentrated loss. A system of social insurance would be much more coherent as it allows every defendant to obtain compensation, not just the ones that were lucky enough to be hit by deep pocketed injurers. Even if compensating every victim would result in only partial compensation- holding constant the total amount of compensation paid by the class of injurers- for the very principle of decreasing marginal utility, social insurance should be favored by the same scholars who claim that tort law has to achieve loss-spreading (Note 6).

The opposite axioms on which the two theories rests, and the embracement by economists of exogenous goals, lead Weinrib to write that “This difference does not preclude the two approaches from arriving at the same results. [however] concurrent results would not efface the theoretical differences that generated them. Nor, of course, would these results indicate the existence of a mixed theory. All we would have is a coincidence of results from two independent theories” (Weinrib, 2002 p. 628). In other words, deterrence would be an effect of the law, instead of being a cause.

The door is opened to the possibility that corrective justice and deterrence theory, while incompatible in the way presented above, may still coexist as conceptually sequenced ideas. In the words of Weinrib “In this sequenced argument, corrective justice is prior to deterrence because it illuminates the nature of the wrongs that positive law deters. Deterrence is then necessary as a further element in this sequence by virtue of being implicated in the actualization of corrective justice through the legal institutions of positive law (Weinrib, 2002 p. 639).

Clearly this idea will hardly please deterrence theorists as they try to show how tort norms should be inspired by deterrence, while a sequenced argument denies that deterrence has any relevance for the content of the norms themselves. As Weinrib (2002) himself admits “situating deterrence within a conceptually ordered sequence that
includes corrective justice affirms both corrective justice and deterrence without resolving the tension between them when each is claimed a ground of the norms” (Weinrib, 2002, p. 639).

5. The Necessity of Corrective Justice to Explain Tort Law

Being an economist, I have always considered deterrence as being the obvious *raison d'être* of tort law. To read the writings of Aristotle and his modern disciples eroded this dogma. To build a legal system on the theory of deterrence, as developed by the most intransigent economists, presupposes a series of assumptions that it is hard to feel comfortable with. To mention a few: agents would have to be able to estimate the probability of every accident, the probability that their conducts are discovered, and that they are actually sued by the counterpart; furthermore they need to be perfectly aware of the legal standards (that by the way needs to be optimal as well), they need to know the magnitude of the fine, or the amount of damages attached to a certain conduct, and so on. While achieving a good level of deterrence still intuitively seems a desirable goal, the idea of optimal deterrence appears to be so far from reality that it should be considered little more than a chimera.

Furthermore, the fact that every legal system takes very seriously the bilateral nature of the relationship between the doer and the sufferer can hardly be explained as a mere coincidence, or as the cheapest way to achieve optimal deterrence.

No matter how rooted is economic thought into oneself, it is very hard to claim that corrective justice does not offer a powerful explanation for the existing tort law. Most of the features of modern tort law seem to be perfectly explainable from a corrective justice standpoint, and such a theory provides a straightforward focal point around which policies can be shaped. At the same time, the line of reasoning that leads Weinrib to claim that corrective justice and deterrence simply cannot contemporarily determine the contents of the norms seems to be correct.

Nevertheless, it appears that there is still a piece of the puzzle missing.

Corrective justice rests on two very strong assumptions: (i) the doer is identified and is actually held liable. There can be no corrective justice without compensation. In other words, as stated in the introduction, for corrective justice to be achieved the probability that a certain conduct is discovered should tend to 1. (ii) The wrong does not destroy resources, only transfers them from the sufferer to the doer. Aristotle explicitly writes “Let the lines AA’, BB’ and CC’ be equal to one another; from the line AA’ let the segment AE have been subtracted, and to the line CC’ let the segment Cd have been added, so that the whole line DCC’ exceeds the line AE’ by the segment CD and the segment CF; therefore it exceeds the line BB’ by the segment CD”. (Aristotle). This description clearly underlines how there is merely a transfer of resources, not what is known in economics as a deadweight loss.

Let us recall now the equation (2) and let us remember the conditions imposed by (3) and (4):

\[ P_f* M_f + P_d*D = H + DWL \]  

Since the portion of the segment is merely transferred from one party to the other, DWL=0. Moreover, as we have stated above, in order to achieve corrective justice the doer has to be held liable, and damages have to be awarded, thus \( P_d=1 \). Lastly, the compensation received has to equal the harm suffered. The equation (2) simply becomes \( D=H \)

This is merely a different way to restate Aristotle’s line of reasoning in the passage cited above. EA is equal to CD, and should be taken away from the doer and assigned again to the sufferer.

What should be underlined is that under these assumptions also optimal deterrence is achieved. If the interaction between the doer and the sufferer takes place in an ideal world, corrective justice cannot be attained without achieving also optimal deterrence (Note 7).

It is obvious that, apart from some really odd circumstances, none of the assumptions described here would hold, nevertheless it is important to be aware that in an ideal setting the two doctrines are perfectly compatible, as they necessarily lead to an identical outcome.

It should be noted that while perfect corrective justice automatically implies optimal deterrence, the reversal does not hold. In fact, in an ideal world, optimal deterrence can be achieved for any suitable combination of \( P_d \) and \( D \), while corrective justice requires \( P_d \) to be equal (or at least approaching) to one. The fact that calibrating the relevant parameters in order to achieve corrective justice, automatically implies that also optimal deterrence is achieved while the opposite does not hold, seems to suggest that the idea of conceptually sequenced arguments is effectively the best compromise available.

From this perspective, the conclusions that should be drawn are not far from the ones drawn by Weinrib: norms
should be grounded solely on corrective justice considerations, and deterrence only comes as a consequence of the norms and institution created in order to achieve corrective justice (Weinrib, 2002).

Even though this line of reasoning appears to be flawless, it points to the first rift in the castle erected by corrective justice theorist: social institutions that enforce the law are far from perfect, and come at a cost that is ultimately borne by each doer and each sufferer. In the next section the main problems created by an approach based exclusively on corrective justice will be examined.

6. Shaping the Target

6.1 The Limits of Corrective Justice

It is hardly possible to find anything written on corrective justice that does not underline the crucial link between the doer and the sufferer on which the whole law of torts is based.

Although the importance of this link is not being questioned, there is something misguiding in the way it is usually worded. In fact, it is generally stated that such a link is created by the wrong, while it appears to be more correct to state that the wrong is part of the link.

A slightly modified version of an example offered by Aristotle can enlighten the difference: Let A be a builder, B a shoemaker, C a house and D a shoe. If A and B decide to exchange their products C and D, a link between the two is established. In the eyes of the Philosopher the bargain should be characterized by a proportionate requital, only then the city can be held together.

In fact, considerations of rectificatory justice arise in connection with transactions both voluntary and involuntary (Aristotle). In both cases the predominant trait is the link between the parties, and in both cases the goal is to achieve corrective justice.

If there are no relevant constraints, the builder and the shoemaker will exchange their products only if neither of the two will be worse off; therefore the bargain will take place only if they obtain something that they consider to be worth at least as much as the price they are paying (Note 8).

For obvious reasons Aristotle’s idea of exchange does not coincide with the one usually employed in modern economics, yet the two do not appear to be mutually exclusive. In fact he writes “[that] demand holds things together as a single unit is shown by the fact that when men do not need one another, i.e. when neither needs the other or one does not need the other, they do not exchange” (Aristotle). There is clearly no mention of Pareto efficiency, surplus from the trade, and all the terminology that is present on any modern economic textbook, yet the underlying idea seems to be very similar.

Furthermore, Aristotle explicitly states that “to have more than one’s own is called gaining, and to have less than one’s original share is called losing, e.g. in buying and selling and in all other matters in which the law has left people free to make their own terms” (Aristotle).

In other words, it seems to be firmly rooted in Aristotle the idea that parties engaging in voluntary transactions can set the terms they consider to be better for them. The positive opinion that Aristotle has of voluntary trade, might induce one to think that parties who willingly engage in a transaction should generally opt for terms that imply a proportionate requital, and that will keep the city united. Given the importance that Aristotle attributes to the poleis, to preserve voluntary transactions can be considered a priority. This conclusion is perfectly compatible with the principle of modern economics that any voluntary transaction leads to an increase in total welfare.

Now let us assume there are only two builders in our polis A and E, and that all the shoemakers’ need a house. Let us also assume that A and E stipulate a secret agreement to artificially raise the price of C in terms of D for all the shoemakers. The shoemakers are now forced to accept the terms imposed by A and E, and to pay an excessive price. These transactions are in the domain of what Aristotle calls involuntary transactions, since the shoemakers are not free to make their own term.

This is the stage where corrective justice comes into play if we consider the link to be created by the “wrong”. Theoretically the builders should be forced to return the surcharge to the shoemakers that have bought a house. Moreover, as suggested by basic economic theory, in order to raise the price, the builders had to restrict the output, hence some of the shoemakers that would have bought a house at a “fair” price, had to sleep on the street and also need to be compensated.

The problem is that perfect corrective justice simply cannot be achieved. It is simply impossible for builders and shoemakers to return to the situation before the overcharge was applied. In fact to ask the builders to simply return the extra profits they made is far from being sufficient.
Some shoemakers were forced to sleep on the street due to a wrongful conduct. How should corrective justice be achieved? If A and E are simply asked to produce more houses and sell them to the homeless shoemakers this would not suffice to repay them for the nights they spent on the street. Conversely if buildings will be forced to pay damages in addition, their situation would become worse than the initial one. In the modern world it is very hard to imagine that A and E could be forced to build more houses, it is much more likely that they will be asked to compensate all the shoemakers in monetary terms; it follows that if we consider as a loss only the surcharge paid by the shoemakers who bought an apartment, then the ones who did not get a house will not be properly compensated. The most likely outcome is that A and E will be asked to refund every shoemaker in accordance with the loss they suffered; such a compensation outweighs the gains that were made, hence the shoemaker would not be brought back to the situation in which they were before the interaction.

To use Aristotle terminology, it is impossible to satisfy the condition imposed by (7), simply because the sum of the segments is not the same before and after the interaction.

This conclusion can be generalized with regards to antitrust infringements since (at least at a theoretical level) any antitrust violation generates a dead weight loss, and hence, reduces the total length of at least some of the couples of segments. In other words, it will never be possible to completely undo the wrong, and to bring all the pairs of doers and sufferers to their original conditions.

This point becomes obvious if we consider other accidents that destroy resources in a more evident way. Product liability or car accidents are prominent examples. A firm has very little to gain from a product that explodes in the hands of its customer, causing a serious injury. It might be argued that the gain could be represented by the savings in precaution costs, yet it is a heroic claim that there is an equality of any sort between them and the harm, and especially that a monetary compensation can recreate this equality.

From this perspective it appears that it makes very little sense to shape the law so to achieve a goal that is simply beyond our reach. It would be much more interesting to develop a comprehensive theory to understand which deviations from the ideal corrective justice should be accepted.

It would not be wise to embrace such a conclusion precipitately, since the friction between the real world and corrective justice as developed by Aristotle seems to be little more than an illusion.

Aristotle idea appears to be that voluntary trades create a binary link that is as strong as the one usually underlined by corrective justice theorists with regards to wrongs. The Philosopher is suggesting that both voluntary and involuntary interaction create a binary relationship between the parties, which embraces the whole interaction between them, not just the wrong. To use the abovementioned example, the link between the consumer and the firm is created when the defective product is purchased (voluntary interaction), not just when the product explodes (involuntary interaction) causing the harm.

In the first part of the binary relationship between the firm and its consumer, the justice is achieved through proportionate requital. When the product explodes the binary relationship between the parties is altered and the judge tries to equalize it. Conversely, since a voluntary relationship appears to be way more desirable, not only for the society (polis), but for the parties themselves, there is no reason not to create norms that will prevent voluntary relationships to be altered by involuntary interactions.

As it has been stated at the beginning of this section, it is not being contested that corrective justice is the inescapable starting point if we want to understand tort law, nor it is being claimed that deterrence and corrective justice should simultaneously coexist in determining the content of the law. The idea of a conceptually sequenced argument that moves from corrective justice as the one presented by Weinrib has been completely embraced.

However, instead of claiming that corrective justice and deterrence are lying on a straight line, it seems that they might be lying on a circle. Norms should be shaped to guarantee that corrective justice is ensured (both voluntary and involuntary acts), but, since voluntary interactions are preferable, and involuntary ones cannot always be entirely corrected, norms should be created to avoid that voluntary interactions are altered by wrongs. In turn, these norms will affect the way wrongs should be treated when they cannot be avoided. A clear example is constituted by precautionary measures, as they interfere with parties’ interaction before the wrong actually takes place.

Not only corrective justice and deterrence should inform the norms, even though via a sequentially ordered argument, it seems that the one without the other is detached from the reality.

Deterrence theory cannot illuminate the connection between the parties, and without the cage of corrective justice it leads to embrace exogenous goals that are heterogeneous and partially incompatible with the core
characteristics of tort law. Corrective justice can illuminate how to handle voluntary and involuntary transactions; however it cannot take into account that the former are a value that should be preserved.

6.2 Equality of What?

Up to this point a very obvious question has been intentionally avoided. Parties should be equal in terms of what? How can we restore equality if we do not know the object of this equality?

The reason why this question has not been answered is simple: it is totally irrelevant for our inquiry. No matter what we use as the relevant variable to measure equality, the argument presented here still holds.

To prove this point we will first assume that the equality is defined in terms of Kantian rights and then in purely monetary terms.

Understood as a manifestation of Kantian’s rights, private law protects rights, not welfare (Weinrib, 1995, p.141). According to Kant, rights are of two kinds: the right to bodily integrity and the right to external objects of the will (Kant, 1797).

Moreover, the relevant gains and losses are normative not factual. In fact, if a firm induces one of her competitor’s clients to opt for her products, it is surely causing a factual loss to its rival, yet such a loss is of no interest for private law as it is not normative in character.

Let us go back to our example of defective product. In the framework developed by Kant the firm realized a normative gain, but no factual gain, while its customer suffered both a factual loss and a normative loss.

Coherently with the idea of justice developed by the German Philosopher, the focus is on the correlative normative gains and losses of the parties. Specifically, the normative loss of the victim consists in the violation of her right to bodily integrity, while the normative gain for the firm results from negligently injuring her.

The only means available to the court to correct the normative loss suffered by the plaintiff is to award monetary damages. However, to claim that a certain amount of money is enough to perfectly restore her right to bodily integrity appears to be very optimistic, to say the least.

Using economic terminology, it is easy to see that money cannot be considered as a perfect substitute of the right violated, otherwise bodily integrity would simply be tradable. To put it more simply, money will not buy the plaintiff a new hand, nor is it possible to define exactly the value of the one he has lost (or of the right of bodily integrity that was violated) as it is proven by the enormous variance in the compensations awarded for serious body injuries.

In other words, due to the fact that compensation can only happen through money, and that money cannot be considered a perfect replacement for Kantian rights, the sum of the length of the two segments will inevitably be shorter after the product explodes, and perfect corrective justice cannot be achieved. The only way to achieve equality is to preserve the voluntary transaction, trying to avoid that it is altered by the involuntary interaction.

It is straightforward that this line of reasoning applies to any possible variable that is not a perfect substitute of money.

Conversely, if we decide to define equality in terms of a variable that can be perfectly replaced by money or in terms of wealth itself, we have to conclude that an interaction that destroys economic resources inevitably shortens the sum of the segments making corrective justice impossible to achieve.

Therefore we can conclude that for any choice of the relevant variable that defines the equality between the parties, an involuntary transaction can shorten the size of the segments making perfect corrective justice unattainable.

6.3 Differences between the Two Cases

The reader has probably noticed a sharp difference between the two examples sketched above. If we accept that the tort is only a part of the interaction between the parties (and that corrective justice and deterrence are in the circular relationship proposed here), it is easy to fit product liability into the framework that we have developed. The link is established when the consumer buys the product (Note 9), therefore a satisfying voluntary interaction is realized. When the product explodes, there is an involuntary circumstance that alters the balance between the parties, which are no longer equal. However, the firm cannot completely rectify the wrong (it has no extra hand that could replace the one lost by the victim in the explosion) and hence the wrong cannot be entirely undone.

Private law should support the creation of a link between the parties, and handle the consequences of the involuntary interaction as suggested by corrective justice considerations. Nevertheless, whenever an involuntary interaction is prevented altogether, the equality between the parties remains intact and a more satisfying result is
achieved. It should be noted that there was no need to introduce nebulous concepts like total welfare, to assign exogenous goals – like loss spreading- to private law, or to treat it as public law in disguise (Green, 1959). The attention is exclusively on the parties and how a just interaction between them is achieved.

Conversely, in the case of the cartel formed by A and E the situation appears to be hazier. A link is certainly established between each shoemaker that actually purchases a house, and the builder that sells it to him. For the reasons stated above, these transactions can be considered involuntary, and corrective justice is achieved if the surcharge is returned to the shoemaker. However, the quantity of houses available is lower due to the existence of the cartel, thus even after the surcharge is returned; there are still some shoemakers that have not been compensated: the ones that were not able to buy a house as a consequence of the limited supply. How can we claim that there is a bilateral link between the builder and the shoemaker if no interaction of any kind has occurred? The only way to answer this question is to introduce the concept of potential customers, yet this appears to be a very unsatisfactory path from the corrective justice perspective.

This might induce one to think that the circular argument cannot be adapted to every interaction regulated by private law. On the contrary it should prove exactly the opposite. Antitrust law can hardly be considered as a purely private body of laws, and it is exactly the fact that it stands in the middle between public and private law that creates friction with the framework developed here. It is simply impossible to explain antitrust law exclusively in terms of private interaction, ignoring the economic consequences of agents’ behavior.

Especially when a goal like protecting consumer welfare (or any other mixed goal) is introduced into the picture, antitrust regulation becomes incompatible with Aristotle teachings, as it clearly mixes corrective justice and distributive justice.

Due to the mixed nature of antitrust law it is hardly imaginable to rely solely on an approach that is typical of private law like the one developed by Aristotle.

7. The “Second Generation” Mixed Theories: Commonalities

Following Schwartz’s example new attempts have been made at creating a mixed theory that could accommodate corrective justice and deterrence. From this perspective, the theories advanced by Geistfeld and Chapman call for special attention and will be further investigated in this section. Even if both advocate the idea that a mixed theory is not only desirable but it is also inevitable, it is important to explore commonalities and differences with the thesis advanced in this work.

The most obvious point of contact between the framework developed here and the theories advanced by Geistfeld (2001) and Chapman (2001) is the idea that corrective justice and deterrence not only can coexist, but that they ought to (Geistfeld, 2009).

From this perspective, Chapman (2001) (building on the sophisticated tools of social choice) takes a step further and tries to suggest concrete ways to accommodate the two allegedly heterogeneous goals. More importantly Chapman suggests the adoption of a conceptually sequenced argument, thus it seems that no relevant friction exists with the framework developed here.

Similarly Geistfeld (2009) recognizes the “symbiotic relationship” between economic analysis and normative principles, demonstrating how the former is often needed to give a practical and concrete meaning to the latter. It is also correctly underlined how the need to incorporate deterrence concerns is strengthened by the imperfect compensatory mechanisms that are available to the courts.

7.1 The “Second Generation” Mixed Theories: Differences

Moving from the idea of path dependency, Geistfeld (2001) tries to address the most fundamental objection raised by moral philosophers: the structure of tort law system hardly seems the optimal choice to minimize accident costs, thus it cannot be considered coherent with the goal of optimal deterrence. It would be very puzzling, or so they say, to pursue a given goal by creating something that is inappropriate for the task.

The answer provided by Geistfeld (2001) can be divided in two parts: (i) behind tort law structure there are historical reasons that justify its existence; even admitting that an omniscient legislator concerned with efficiency might adopt a different solution if he had to craft a legal system starting from zero. (ii) It is possible to offer a possible interpretation of tort law in terms of economic efficiency.

With regards to the first claim Geistfeld argues that “the tort system was initially designed in the twelfth and thirteenth centuries to implement corrective justice for cases in which the defendant criminally injured the plaintiff...[however] legal decision makers rejected natural-law justifications in favor of more pragmatic,
instrumentalist justifications… The tort system could not wholly reject its corrective origins in favor of an overtly instrumentalist approach, as any change in judicial decision-making is constrained by the requirements of *stare decisis*” (Geistfeld, 2001, p. 254).

While his historical description is certainly accurate, a few caveats should be made. Firstly, as Geistfeld admits, the fact that some regulatory purposes were introduced into the picture does not automatically lead to accidents costs minimization (Geistfeld, 2001). Secondly, to assign such a central role to *stare decisis* appears to be in sharp contrast with the evidence from legal systems in which precedents are not binding, as they are closer to corrective justice than common law countries. It is no secret that in European countries corrective justice occupies a more prominent position than in US and in UK. In other words, either civil law countries have developed equally effective means to preserve the coherence of their legal systems and the uniformity of the law over time, or there had to be another reason not to abandon corrective justice. It should be noted that if the former explanation is accepted, the enormous body of economic literature that underlines the wonders of *stare decisis* would lose much of its credibility (Macey, 1989). However, even if we admit that the distinguishing role of *stare decisis* has been systematically overstated (Mattei & Pardolesi, 1991), we are still left without an explanation for the persistence of corrective justice over the centuries. An alternative justification could be introduced abandoning the extremely reductionist approach typical of welfare economic, and allowing holism to play some role. In order to introduce this justification it is necessary to sketch briefly the centuries old debate between holists and reductionists.

7.1.1 Reductionism

There is no univocal definition of methodological individualism (Lukes, 1969; O’Neill, 1973; Udehn, 2001 and 2002) and the crucial ambivalence is whether the *explanantia* of social phenomena has to be found in individuals alone or in individuals plus relations between individuals (Hodgson, 2007). According to Popper (1945), “methodological individualism rightly insists that the “behavior” and the “actions” of collectives must be reduced to the behavior of the actions of human individuals. The use of the term “reduced” implies that the goal of social science should consist exclusively of statements about individuals (Hodgson, 2007).

Societies can be considered as an aggregation of individuals, and hence welfare maximizations (or accidents costs minimizations) can be achieved through the maximizations of private welfare aggregated according to the relevant criteria.

7.1.2 Holism

At the other end of the spectrum, proposers of what Phillips calls Holism 2, advocate the five theses of organismism: “(i) The analytic approach as typified by the physico-chemical sciences proves inadequate when applied to certain cases …. (ii) the whole is more of the sum of its parts, (iii) the whole determines the nature of its parts, (iv) the parts cannot be understood if considered in isolation from the whole, (v) the parts are dynamically interrelated or interdependent” (Phillips, 1976, p. 6) plus the additional idea that a whole, even after it is studied, cannot be explained in terms of its parts. From the perspective of an organist individuals are only a part of the whole society, thus it is not possible to gain any understanding of the individuals if we do not have knowledge about society as whole.

The main claim here is that merely aggregating individuals’ welfare does not provide exhaustive information on the welfare of a society considered as a whole.

An economist will probably perceive this claim as absurd given the dominance of welfare economics, while some social scientists might even consider it dangerous due to the gross misinterpretation that totalitarian states made of Hegel’s (1807) and Gierke’s (von Gierke, 1902) theories respectively on ethical state and on human groups.

However holistic thought is a fundamental component of most sciences, and beside historical reasons there seems to be little justification to rule out a priori its relevance in a context where wholes (human societies) are extremely complex (Note 10).

To admit the possibility that social welfare is not a mere aggregation of individuals’ welfare opens the door to the hypothesis that a concept of justice embedded in the society has an independent value. On the one hand, this would explain why corrective justice has not been abandoned by courts in any country, regardless of the existence of *stare decisis*. On the other hand, at a broader level it explains why, in spite of the attempts made by Kaplow and Shavell (2009) to offer an “anti-darwinist” (Note 11) explanation of legal systems evolutions, the concept of justice has been present in a form or another in every human society.

The third and more fundamental concern raised by this historical explanation regards the causes underlying the
“lock-in” effects created by tort law. (Note 12) The more plausible explanation offered by Geistfeld is that “even if it would be cost-effective to change regulatory institutions, that change must be legislatively implemented. The various costs that individuals or groups would incur to the displacement of the tort system give them a substantial incentive for forming interests groups to defeat such legislation.” (Geistfeld, 2001, p. 257) However “interest groups are usually well organized by people with strong gains from particular government actions” (Romano & Thammapitakgl, 2013, p.67), and hence, to introduce the possibility that the macro-system – tort law as a whole - exists as a response to lobbying, implies that also the single norms composing the macro system were implemented and are defended due to interests groups’ pressure. In fact, there is very little reason to assume that interest groups only engage in defensive lobbying and do not actively promote the implementation of legal rules that might favor them. It is also well established that it is easier to organize pressure to address a specific and narrow issue than to promote a widespread interest like the preservation of the structure tort law as a whole, since the number of parties concerned is bound to be relatively smaller and their interests are likely to be more homogeneous.

However if we accept these hypotheses, very little is left of cost minimization concerns, as tort law is described as a mean to protect the interests of the most powerful and better organized interests groups.

In other words, the historical account described by Geistfeld seems to lead to a mixed theory between corrective justice and the Marxian idea that law is merely a mean to protect dominant classes (or powerful interest groups), more than a compromise between corrective justice and efficiency.

8. Conclusions

“For we are inquiring not in order to know what virtue is, but in order to become good, since otherwise, our inquiry would have been of no use.”

Aristotle, Nicomachean Ethics book II.

For over two millennia corrective justice has been the foundation of tort law, and private law in general. Most instrumentalist approaches that are so much à la mode, pursue goals that are in contrast with the basic features of tort law, and hence appear inadequate to offer a comprehensive account of its characteristics.

No theory has been developed that can replace corrective justice, and no satisfying explanation has yet been offered on how it could coexist with deterrence.

The idea of pure corrective justice is surely very appealing; however it appears to be a utopia, since involuntary interactions tend to irremediably alter the equality between parties and enforcement is far from perfect. As the cited passage proves, it should not be neglected that Aristotle inquiry was practical in nature. He did not aim to develop a framework to achieve justice in an ideal world, but to improve the concrete one he was facing. To disregard that the probability of compensation will very rarely be close to one, or that the complete undoing of a wrong is often a chimera, means to transform Aristotle ethics into something completely abstract.

Corrective justice is fundamental to underline the indissoluble link between the parties, and to offer an account of the main features of tort law; however it completely ignores the fact that voluntary transactions are a value that should be protected. Not to maximize social welfare or to spread losses, but to preserve a just interaction between the parties.

Corrective justice and deterrence should therefore be considered conceptually sequenced ideas; however, they appear to be lying on a circle, instead of a straight line. Not only corrective justice and deterrence should both inform the norms, even though through a sequentially ordered argument, it also seems that the one without the other is detached from the reality.

On the one hand, deterrence theory cannot illuminate the connection between the parties and, without the beneficial influence of corrective justice, it tends to include exogenous goals that are heterogeneous and partially incompatible with the core characteristics of tort law. On the other hand, corrective justice explains how to handle voluntary transactions and involuntary transactions, while it ignores that the former are a value that should be preserved.

Acknowledgements

I am grateful to Pardolesi, Visscher and Figà-Talamanca for their comments, their corrections, and especially for their patience.

References


Notes

Note 1. We are now considering the case where harm is equal to the gain of the tortfeasor. If the harm and the gain are not identical it is controversial which measure should be used and under which circumstances. In any case the point raised here, that \textit{prima facie} there is a clear difference between the goals of corrective justice and deterrence theory, remains valid, no matter which measure is adopted. For an insightful discussion of the problem see Weinrib (2000).

Note 2. The label total liability will be used to denote the sum of the fine and the damages. Symmetrically total expected liability indicates the sum of the expected fine and the expected damages.

Note 3. It is implicitly assumed that the goal of antitrust regulation is to maximize global welfare.

Note 4. It might be important to clarify the terminology. The label “Harm” indicates only the harm directly suffered by the victims and that can be compensated. The Deadweight loss includes (i) the resources destroyed by the anticompetitive conduct (ii) the indirect harm caused to victims that has no possibility of realistically be taken into account. For example a cartel might create negative spillovers on the consumers of a contiguous market; however this will most likely not be taken into account in an antitrust case.

Note 5. To be more precise if the victims are consumers H should be defined as the reduction in consumer surplus.

Note 6. From the argument proposed here it should not be inferred that we have to opt for social insurance. The only implication is that it is crucial to investigate the consequences of introducing certain goals into the picture, if we want to develop an effective set of laws. We cannot simply claim that if tort law was introduced in every developed legal system, then it must be the best possible solution to maximize social welfare. If the existence of something implies \textit{ipso facto} its superiority in terms of efficiency, then the utility of law and economics would be minimal.

Note 7. It is important to note that in the ideal world described by economists the administrative costs are assumed to be equal to 0; therefore D=H is a satisfying equilibrium. See Shavell (1980). It is important to note that even if administrative costs are taken into account this solution would still minimize primary and secondary accident costs. The discrepancy between deterrence and corrective justice would be then limited to tertiary costs. For a definition of primary, secondary and tertiary costs Visscher (2009).

Note 8. It should be noticed that Aristotle includes gratitude, friendship etc. in the calculus. This inclusion is
compatible with economic theory.

Note 9. It is obviously possible to imagine situations in which the creation of the link precedes the moment in which the product is bought

Note 10. It suffices to think to the problems we would encounter if we attempted to determine the characteristics of human body (even the most fundamental ones as life) by observing and aggregating the properties of the quanta present in it.

Note 11. With the term “anti-darwinist” here it is intended that for the authors the evolutionary process systematically resulted in the inferior outcome of societies based on fairness concerns. A devastating criticism to this argument on completely different grounds is advanced by Coleman (2003).

Note 12. The importance of his metaphor of the “curvy road” should not be overestimated since it simply proves that some costs are associated to every change.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).