

**WHY LAW BREEDS CYCLES**

**Alvaro Sandroni**

Northwestern University  
Kellogg School of Management  
sandroni@kellogg.northwestern.edu

**Leo Katz**

University of Pennsylvania Law School  
lkatz@law.upenn.edu

## **Abstract**

Intransitive choices are generally held to be the mark of irrationality. When a set of rules engenders such choices it is usually held to be irrational and in need of reform. Many legal arguments involve showing the deficiency of a doctrine, a theory, or an argument, by showing that it leads to intransitivity. In this essay we try to show that rule-based reasoning is by its very nature certain to lead to intransitivity, not just here and there but widely. This raises difficult questions as to what should count as a valid logical objection in legal argumentation. Along the way, it casts new light on the controversy about rights-based legal theories, Kaplow and Shavell's "anti-fairness" theorem, and Sen's liberal paradox.

## WHY LAW BREEDS CYCLES

Alvaro Sandroni and Leo Katz

It is well known that law sometimes gets into logical tangles of the sort where it ranks A above B and B above C, but then, perversely, C above A. The best known of these tangles are probably the problems of circular priorities in commercial law and *renvoi* in conflicts of law. Such problems are generally thought to be mostly in the nature of intellectual curiosities, interesting exotica, more famous than they deserve to be, given their rarity and irrelevance to the real world. But as we will try to show in this essay, the phenomenon they exemplify—cycling—is in fact ubiquitous throughout the law. Really the only thing that makes circular priorities and *renvoi* exceptional is that they are among the few instances of cycling that have been widely noticed.

Cycling, we will try to show, is a nearly inevitable feature of rule-based reasoning, within and outside the law, and we will take a look at it in both contexts. It cannot really be squelched, at least not without producing even more unpalatable difficulties. The connection between cycling and rule-based reasoning has many implications, which we can only begin to probe here. It suggests, for instance, that the intuitively plausible picture many people have of social norms, as being something akin to a preference or a budget constraint is problematic. It also suggests that many criticisms often made of rights-based or deontological approaches to law and morality are probably misguided: the criticisms will turn out to apply to all rule-based reasoning, not merely those with a deontological orientation. Finally, the connection between cycling and rule based reasoning raises a real puzzle about the very nature of legal arguments, which are usually thought to involve a striving for consistency. Without that, how can law act as a meaningful constraint on those who have to apply or live under it? And yet it seems that consistency in the traditional sense of an absence of intransitivity is just not to be had. What now?

## 1. *Some Background*

While our thesis that rules inevitably breed cycles is an unfamiliar one, the Canadian scholar Bruce Chapman has maintained, for many years now, that intransitivity is not a blemish to be avoided, but an essential feature to be explored and better understood. The rationality of the legal decision maker, he has argued, is different from the rationality of the utility-maximizing consumer or producer in the standard economic model. Once the application of rules is involved, we should not expect decision makers to continue to satisfy the rationality constraints of the standard economic models (Chapman 2003).

To make this point, he has repeatedly made use of the following example, adapted from Sen. Suppose, he writes

someone is offered a choice of fruit at the end of a dinner. If only a large apple,  $A$ , and a large orange,  $O$ , are offered to her she should choose the large apple. Both fruits are large and all else equal, she prefers apples to oranges. However, if she is offered  $A$ ,  $O$ , and a small apple,  $a$ , then different considerations arise. For now there is an issue of etiquette to be addressed. The rule, let us say, is that one should never choose the larger of two items of the same kind. Thus our chooser now reasons that, in the choice set  $(A, O, a)$  she cannot now choose  $A$  because that would be in breach of the rule of etiquette. She therefore chooses  $O$ , a piece of fruit that is larger than  $a$ , but a fruit of a different kind. Thus, from the set of alternatives  $(A, O)$  she chooses  $A$ ; but from the set of alternatives  $(A, O, a)$  she chooses  $O$ . (Chapman 2003, p. 1178).

To be sure, what we have here is not strictly speaking an intransitivity. But it is a closely related oddity, and it is not difficult to derive an intransitivity as well. As Chapman explains,

The chooser would . . . reveal an intransitive preference ordering if the different fruits were offered to her in pairs. She would choose  $A$  from the pair  $(A, O)$ ,  $O$  from the pair  $(O, a)$  and  $a$  from the pair  $(a, A)$  in violation of transitivity. The reason of course is that the rule of etiquette does not come into play until the third choice, when the big and small apples are presented together. Until that point the chooser can select between the fruits purely according to taste, or according to the different values of the different alternatives, choosing the highest valued one; in other words, she can choose in the way that the [traditional] theory of [rational choice] . . . suggests. But when the two apples are presented together, etiquette becomes an issue between them (Chapman 2003, p. 1178).

In this essay we will explore the basis for this kind of oddity, its ubiquity throughout the law, and some of its implications. (On cycling in law more generally, see also Spitzer (1979), Kornhauser and Sager (1986), and Stearns (1997).)

***2. The Basic Argument: The Rationalization Model of Cherepanov, Feddersen and Sandroni (Cherepanov, Feddersen, and Sandroni (2008)).***

Consider several fairly mundane choice situations from everyday life, taken from the vast psychological literature on somewhat paradoxical-looking everyday choices.

*Situation 1.* Art decides to take time off from work and go home early. However, prior to leaving the office, he is informed that a colleague is a patient in the local hospital and can accept visitors that afternoon. Arthur reconsiders his decision to go home and stays at work instead. While Arthur's decision might momentarily catch us off guard, even on brief reflection, we will not really find anything mysterious about it. If Arthur were to take time off from work, he would feel honor-bound to visit his colleague in the hospital, which he is not in fact keen to do. If he stays at work, he has a satisfactory reason to give—to himself, as well as others—for not visiting (Cherepanov, Feddersen, and Sandroni 2008).

At a slightly more abstract level, though, there is in fact something mysterious about Art's decision, because he seems to be exhibiting an intransitive series of choices. As between work and movie, he would choose to go the movie. As between movie and hospital, he would choose to go to the hospital. But as between hospital and work, he would choose to go to work. In other words, his choices form a cycle and he therefore appears, at least at a formal level, to be behaving irrationally. The sense of mystery and paradox is compounded if we compare his decision to stay at work rather than going home with his decision to go home when going to the hospital is also an option. That starts to sound like the

proverbial restaurant guest who chooses chicken over steak, if those are the only two items on the menu, but chooses steak, if fish happens to be on the menu as well.

*Situation 2.* Given the choice between watching movie X or movie Y, Bert chooses Y. Given the choice between watching movie Y by himself or watching it in the company of a grotesquely disfigured midget, Bert chooses to watch in the company of the midget. On the other hand, when Bert is given the choice between watching movie X by himself or watching movie Y in the company of the midget, he chooses X. Once again, while this might ever so briefly puzzle us, it won't do so for long. Bert obviously likes movie Y better than movie X. He also feels that he shouldn't refuse to watch the movie in the presence of the midget, much as he detests the idea. But he also feels that if he chooses a different movie to watch, he won't have to admit—to himself, or to others—that he was avoiding the midget. Rather he can say he was keen to see X, which, as it happens, he could only do by himself (Snyder, Kleck, Strenta, and Mentzer 1979).

Although all of this is commonsensical when looked at from up close, once again a more abstract perspective makes what is going on here a bit mysterious because Bert is making cyclical choices: He chooses movie Y over movie X, movie X over movie-Y-with-a-midget and movie-Y-with-a-midget over movie-Y. Bert's choices form a cycle, and therefore, formally at least, he appears to behave irrationally. Moreover, the sense of irrationality can once again be heightened if we compare how Bert would choose if his choice is between movie X and movie Y—namely movie Y—with how he would choose if movie-Y-with-a-midget is added to the menu. Now he opts for movie X.

*Situation 3.* If Chris's alma mater asks him for a small donation, she will say yes. She does not like to turn them down. But if they ask her to contribute either a small sum or a larger sum, she will contribute nothing. While such behavior will seem neither unfamiliar nor irrational from a commonsense perspective, it obviously does seem irrational at a purely formal level. Formally, she too is

acting like the restaurant guest choosing between steak, chicken and fish (Berger and Smith 1997).

One of us—Sandroni, along with Cherepanov and Feddersen—has constructed a formal model to explain what is going on in such cases. The model takes as its point of departure the familiar model of a decision maker maximizing his preferences subject to a feasibility constraint, and adds the following modification: Think of the decision maker as maximizing subject to several constraints at once. The first is the familiar feasibility (or “budget”) constraint. The second is that all decisions must satisfy a set of what we call “rationalization constraints.” Simply put, it must be justifiable under one or another of a limited set of rationalizations. Formally, a rationalization is simply a ranking of some subset of the alternatives the decision maker might encounter, usually reflecting some intuitively plausible ethical principle. In other words, a decision-maker only makes decisions that are both feasible and justifiable. Subject to that two-fold requirement he gets to choose what he likes best. If he faces a bundle of alternatives A, B and C, out of which he likes C best, he can only choose it if there is some rationalization-ranking available that ranks C highest. Otherwise he might have to choose his second-best alternative, let us say A, provided there is some rationalization-ranking that in fact ranks A over B and C. And if there is not, he is stuck choosing his worst alternative, C.

Let’s now see how this model makes sense of the three choice situations with which we began. As to *Situation 1*, involving a series of choices having to do with staying at work, going home, or visiting a colleague in the hospital, we would explain what is going on as follows. Choosing home over work is consistent with the chooser’s preferences and is also consistent with the rationalization that allows one to choose to go home unless work is too pressing. (In other words, there is a particular rank ordering that includes the work alternative and the home alternative, and that ranks home, under these circumstances, ahead of work.) When it comes to deciding between staying home or going to the hospital, however, things are a little different. The chooser would prefer the home alternative, but this is not a choice he can “rationalize.” None of the available rationalization-rankings allows him to do so. Put

differently, there is no way he can justify *not* going to the hospital instead of staying at home. Hence he goes. Now as between visiting the hospital or staying at work, the chooser would prefer staying at work. This is a choice he *can* rationalize. And so that is what he would do. In the same vein, when choosing between all three alternatives, work, home, or hospital, the work alternative is one he can rationalize, but the home alternative is not, which is why among those three he cannot get his first choice either.

The explanation of *situation 2* proceeds perfectly analogously. Given the choice between movie X and movie Y, the chooser can choose Y because he likes it better, and it is presumably rationalizable. Given the choice between watching movie Y with or without a midget, the chooser would prefer the latter, but cannot find a rationalization that would allow him to do so; hence he is stuck with the former: watching Y in the presence of the midget. Finally, given the choice between watching Y in the presence of the midget or watching X, he both prefers X and has no trouble finding a rationalization, inasmuch as he is no longer engaged in patent discrimination against midgets.

For purposes of what follows, this sketchy outline of the model will suffice. That is not to say that there aren't some real gains to be gotten from the formalization. The formal model has a number of implications that are not so easy to arrive at without it—just not implications we need to worry about in connection with the arguments we shall be developing here.

### ***3. An Implication Concerning Social Norms***

The model has an immediate and, it seems to us, noteworthy implication with regard to the concept of social norms, which has received so much discussion only recently. Legal economists sometimes think of social norms in one of two ways, both plausible, and yet, if our model is right, both erroneous, namely as a preference, to be included somehow among other more traditional preferences for certain commodities, or as a feasibility (or “budget”) constraint, somehow part of the traditional way in which the feasible set is conceived, within which the consumer is thought to select his preferred

commodity bundle. Neither view would seem to be right, if one thinks of social norms as being a particular kind of rationalization. For neither preferences nor budget constraints ordinarily lead to cycling, and yet rationalizations, as we just saw, quite clearly do (see for instance, Kaplow and Shavell 2002; Cooter and Porat 2001); on the legal norms literature more generally see Posner 2002).

#### ***4. Legal Counterparts***

If the rationalization model gets it right, then what gives rise to cycling in everyday situations is the fact that someone is maximizing his preferences subject to certain quasi-ethical constraints. It stands to reason that someone who is maximizing his preferences subject to legal constraints should also be producing cycles. And that is just what one finds. What follows are some particularly telling examples.

The first example has to do with the defense of duress, which, following the Model Penal Code's definition, a defendant is allowed to plead when he has committed a crime because he was threatened with serious harm by the person ordering him to commit the crime, unless he complied—provided the threat is of the kind a reasonable person would have found very difficult to resist. More precisely, it requires that the “actor engaged in the conduct charged to constitute an offense because he was coerced to do so by the use of, or a threat to use, unlawful force against his person or the person of another, which a person of reasonable firmness in this situation would have been unable to resist.” (American Law Institute 1981, §2.09).

Now consider the following case. Defendant is threatened with painful consequences—not outright death, let us say, but something like suspension over a fire, with a distinct risk of dying as a result of burns or smoke inhalation—unless he helps the people who have made this threat (and are powerful enough to actually carry it out, sooner or later). Under the circumstances, he yields, and unsurprisingly, is acquitted on grounds of duress--nothing remarkable in any of this.

But now suppose further: Defendant owns a single copy of a manuscript that he has spent the better part of his life writing: his life's work, as he sees it. So important is this manuscript to him that when sometime in the past a fire threatened his house, defendant was willing to incur significant smoke inhalation, burns, and other serious physical harm to save it from destruction.

Finally, consider a further aspect of the law of duress. Imagine a situation in which someone tries to get our self-same defendant to commit a crime by threatening to destroy that valuable manuscript he has spent his life on. Let us suppose that the defendant complies. Does he still get the duress defense? Almost certainly not, since destruction of a valued manuscript is not "unlawful force against his person or the person of another."

Were we to observe the defendant as he encounters these various binary choices seriatim, we would find him in a cycle. He would choose to commit a crime over suffering smoke inhalation, he would choose smoke inhalation over the loss of his manuscript, and he would choose the loss of his manuscript over the commission of a crime. This at least is what we would see if he is law-abiding. The reason is captured by the rationalization model: He would choose smoke inhalation over the loss of his manuscript because that is what he prefers and he has no trouble squaring his actions with the law: the autonomy we all enjoy to take certain risks for whatever it is we sufficiently value. He would choose to commit a crime rather than suffer smoke inhalation because that is what he prefers and he can square this too with the law: namely the defense of duress. But finally, he chooses to sacrifice his manuscript rather than commit a crime, not because he prefers it, but because he cannot square that opposing choice with the law. A manuscript is not a serious enough interest to warrant the commission of a serious crime.

Here is a second illustration, this one involving the defense of necessity. It is similar in structure and flavor but interestingly different in content. The defendant can avoid the risk of starvation to a

group of mountain hikers if he is willing to break into a mountain cabin and help himself to its supplies. Now suppose the hikers would all be willing to endure a high risk of starvation to climb the mountain they are on. That is enough to yield another cycle. We might observe them choosing to risk starvation rather than not climb the mountain. We would also observe them break into a mountain cabin rather than starve. And finally, we would observe them not climbing the mountain, if the only way they could do so would be to break into the mountain cabin. Once again we get a cycle. (American Law Institute 1981, §3.02).

The defendants will choose the risk of starvation over not climbing the mountain, because that is what they prefer and what their right to autonomy allows them to do. They will choose breaking into the cabin over facing the risk of starvation because that too is what they prefer to do and what the law of necessity allows them to do. (“Conduct which the actor believes to be necessary to avoid a harm or evil to himself or to another is justifiable, provided that the harm or evil sought to be avoided by such conduct is greater than that sought to be prevented by the law defining the offense charged.”) But they will not choose to break into the mountain cabin to obtain supplies that would enable them to climb the mountain without too much risk of starvation, because although that is what they would prefer to do, doing so would not avoid harm “greater than that sought to be prevented by the law defining the offense charged,” i.e. burglary and theft.

For a third example of cycling, let us consider the law of self-defense. To avoid getting seriously injured from someone’s attack on him the defendant is allowed to seriously injure him in turn. (“[U]se of force upon or toward another person is justifiable when the actor believes that such force is immediately necessary for the purpose of protecting himself against the use of unlawful force by such other person on the present occasion” (American Law Institute 1981, §3.04). Now suppose that he would be willing to incur serious injury to protect his manuscript from great harm. Once again we get a

cycle: We would observe the defendant when choosing between getting injured or suffering damage to his manuscript, choosing to get injured instead. We would also observe him, when choosing between suffering damage to his manuscript or inflicting harm against the person who is damaging it, letting the manuscript be damaged. Finally, we would observe him when choosing to suffer personal injury or injure his attacker instead, choosing to injure his attacker.

This cycle too fits comfortably within the rationalization approach. As between letting the manuscript be damaged or suffering injury to himself, personal autonomy permits him to choose the latter, and he will be maximizing his preferences and thus able to reconcile it with the law. As between suffering injury to himself and injuring his attacker instead, he will do the latter, being able both to maximize his preferences and rationalize what he is doing under the law of self-defense. Lastly, as between injuring an attacker and saving his manuscript, he will choose to sacrifice his manuscript, assuming he is law abiding, not because that is what he prefers but because he cannot rationalize the contrary behavior.

For a fourth example of cycling, let's look at the law of negligence. Negligence is generally understood to be the unjustifiable imposition of risk. (Criteria for justifiability vary. A commonly invoked one is the Hand formula: Does the benefit of taking a precaution exceed its cost). X and Y are the innocent victims of a traffic accident. To avoid suffering permanent injuries as a result, they speed to the nearest emergency room, in the process imposing a risk in the amount of  $pL$  (probability of loss times amount of loss) on the surrounding world. But if they had not done so, they would have suffered a risk in the amount of  $p'L'$ , much higher let us say than the one they imposed on others by speeding to the ER. Finally, let us suppose they are avid mountain climbers who are prepared to run risks well in excess of  $p'L'$  to get to a mountain top. This yields a familiar cycle. Given the choice between avoiding a risk in the amount of  $p'L'$  and getting to the mountain top, they choose the latter. Given the choice between

getting to the mountain top and avoiding imposing a risk of  $pL$  on others, they are expected to, and therefore will do the latter—not climb the mountain, and avoid imposing risk of  $pL$  on others. Given the choice between avoiding imposing  $pL$  on others, and experiencing  $p'L'$  themselves, they would choose the latter--which once again yields a cycle.

It helps to appreciate the ubiquity of this problem if we realize that it is bound occur whenever there is anything like a hierarchical rule system. By that we mean: there is a constitution which leaves legislators free to act, provided they are consistent with the constitution; which allows judges to act provided they are consistent with what legislators want; and which allows citizens to do what they want provided they are consistent with what judges want them to do; or so one might put it in a highly schematized, simplified way. The result is that we get cycles exactly like the above. So for instance, it might happen that the legislature is free to pass from status quo to passage of Law 1; that it is also free to move from Law 1 to Law 2; but not to pass from Law 2 to status quo. What would be examples of that? Here is a very familiar one. The government is not allowed to punish or tax people for exercising certain rights. It is allowed to impose taxes. It is also, frequently, allowed to grant deductions or other tax benefits, to compensate people for not exercising certain rights they do have. That of course means that they can do indirectly what they can't do directly: by imposing taxes and waiving them if the person does not exercise certain rights they are doing the functional equivalent of taxing them for the exercise of their rights—exactly the kind of cycle which is guaranteed to exist in a regime in which the government is permitted to maximize certain preferences provided it can “rationalize” them.

Although ubiquitous, many, though certainly not all, of these kinds of cycles have gone largely unnoticed--not entirely but largely, at least when compared to more exotic, but much more written about cycles like circular priorities or *renvoi*. The reason actually is relatively simple, but still worth noting. *Our* cycles do not produce dilemmas in which a judge simply does not know what to do. They

might produce what look like inconsistent choices over time, but they don't produce judges who find themselves facing contradictory, because cyclically related, demands within one and the same case. In other words, in the context of circular priorities, a judge faces a rule that tells him that as between three claimants here and now before him, he should favor claimant A over claimant B, claimant B over C, and C over A. In our cases, it is simply the possibility that case 1, he might be called upon to favor course of action A over course of action B, in case 2 he might be called on to favor course of action B over C, and in case 3, course of action C over A. In actuality, this is unlikely to materialize.

### ***5. Can the Cycles Be Avoided?***

Ordinary budget constraints, and more generally, feasibility constraints, don't lead to cycles. So presumably if somehow one could turn these constraints into something resembling an ordinary budget constraint, the cycles would go away. How might one accomplish that? One way to do that is to adopt something resembling the carbon trading approach. Let's see how that would work—and why ultimately it would not. Let's use as an example the negligence cycle described above. In particular, consider the following scheme which is designed to get rid of the negligence cycle in the context of traffic accidents. Suppose that the government treated risk the way the carbon trading system would treat emissions. It chooses a certain amount of risk—expressed perhaps in the number of expected traffic fatalities—which it would be socially undesirable to exceed. It then allocates to each driver a certain number of “risk-points”, which he is free to spend or sell as he sees fit. As long as the amount of risk-taking traffic is limited to the number of risk-points given out, the number of traffic fatalities stays below the target amount. Under such circumstances, the negligence circle we constructed could not, it seems, arise. That's because the amount of risk I am allowed to impose no longer depends on exactly why I am imposing it—whether it be to climb a mountain, or to survive in the wilderness.

But this approach would of course have a number of fairly uncomfortable implications. To start with the most obvious: If one wants to drive to a lunch appointment in the same risky manner that would be appropriate for getting a group of desperately injured accident victims to the emergency room one is now free to do so, provided one is willing to spend a large enough portion of one's point total on it.

Only slightly less obvious, there will be the difficulties of fixing the appropriate risk total to correspond to the total amount of traffic fatalities. This will be rendered even more complicated by the fact that risk is something that is highly information-relative. What seems risky today, given what we know today, won't be deemed so tomorrow with a bit more information at hand. What then is the time as of which the risk someone has imposed is to be calculated? At the moment at which the points are awarded? At the moment at which he starts to act? Partway through the act? Or something else?

Finally, and perhaps most significantly, we will not in fact have done much in the way of alleviating the cycling problem. We have simply moved it to another location. Let us see why. A risk can only be imposed if paid for by points--provided the risk is imposed in the context of negligence. Things are entirely different if the risk is imposed in the context of, say, self-defense, or necessity, or duress. That means not only that the self-defense, duress or necessity based cycles will of course persist. But new cycles will arise as a result of the new framework for dealing with negligence-created risk because risk can now be created on the one hand by paying points in a negligence context and on the other hand, without paying points, outside a negligence context, e.g. a self-defense or duress or necessity context. Let's see how that gives rise to a cycle. Rather than impose risk in negligence context you might benefit from arranging to impose it in a necessity context, where you can do so free of charge, as it were.

#### ***6. Are These Cycles something to Worry About?***

We worry about cycles because transitivity seems to be the hallmark of rationality. If law generates cycles, is law irrational?

Economists like to dramatize the irrationality of intransitivity by appealing to the money pump. If I prefer A to B to C and the latter again to A, it seems I could be made to pay money to someone so as to be moved from A, to B, to C, and thence back to A again. One can certainly, in a purely abstract sense, imagine this being done with our examples. Given that the defendant would prefer smoke inhalation to losing his manuscript; and he will be willing to pay money to get that option. Similarly he would prefer committing a crime to having to suffer smoke inhalation; he will therefore be willing to pay money to get *that* option. And finally, he will feel obliged to sacrifice his manuscript rather than commit a crime; and will thus feel obliged to pay for that last rearrangement as well. The money pump argument presumably has special cachet with economists because it focuses on the kind of consequence they have been trained to be most alert to—losing money through arbitrage. But whether the money pump argument does anything beyond driving home the message of irrationality in language an economist can most readily understand is less clear. To lawyers, the money pump worry may seem even more artificial, because the kind of transactions necessary to effectuate it seems so hard to envision in the context our examples. In particular, it requires a decision-maker to pay in order to abide by the law.

And yet something very much like the money pump worry is a very realistic concern in the law: the possibility of circumvention that inheres in the very existence of such cycles. But there is a way in which the intransitivities presented are quite likely to raise a very practical problem for the law: they are ripe for strategic exploitation. Here is a simple example. Let's go back to the case of duress discussed earlier. This is the case of the defendant who is threatened with something very painful unless he helps the people who have made the threat commit some crime, for which, we noted, he could validly claim the defense of duress. By contrast, we noted, if he had been threatened with the destruction of a

treasured manuscript he has labored over for many years, and if, to avert the manuscript's destruction, he had assisted them in their planned crime, he would not qualify for the defense. This gave rise to a cycle, we observed, because the defendant is allowed to choose to endure great pain in exchange for protecting his manuscript, he is allowed to commit a serious crime so as to avoid being subjected to the painful treatment, but he is not allowed to commit the crime so as to prevent his manuscript from being destroyed. Here is how he might exploit this intransitivity strategically: Suppose the defendant is determined to do the equivalent of saving his manuscript by committing a crime. What he does is to pay off the people who are seeking to recruit him for a crime, with money he borrows from a loan shark. This loan shark in turn demands that he commit a serious crime as a way of extinguishing his debt, which he cannot pay. But if he committed a crime to escape the loan shark's threats, he would most likely qualify for the defense of duress, *because at this point he is doing so not to protect his manuscript but to avert great physical harm*. The example is of course contrived. But the contrivance is the sort that is bound to have many more realistic counterparts. Wherever there is an intransitivity, there is an opportunity for strategic exploitation.

But let us cut to the heart of the matter and ask directly whether the law is being irrational when it permits this kind of cycle to exist?

At first glance, it might seem not. The internal psychology of what is going on here does not seem irrational. It does not feel irrational to say that you cannot commit a crime to protect your manuscript, but when something far weightier is involved, namely the safety of your person, that's a different matter. The same holds for the other examples of cycling.

At second glance, the matter seems no longer so clear at all. Why, for instance, do we not say about the duress case: The defendant was willing to accept smoke inhalation for the sake of the manuscript. We should thus be willing to insist he do the same for something we deem far more

important—avoiding the commission of a crime. Why do we not say something analogous about the necessity case? The defendants are willing to risk starvation for the sake of reaching the mountain top; should we not therefore insist that they do the same for the sake of something we think far more valuable, avoiding a break-in? Why finally do we not say the same thing about the self-defense case? The defendant was willing to risk serious personal injury for the sake of his manuscript. Should we not therefore insist that they run the same risk for something we consider more important, avoiding the killing of another person, albeit an attacker?

There lurks behind all of this an even larger problem. Requirements of coherence are the main constraints on legal reasoning. Critiques of law often depend on finding inconsistencies, for instance in the way in which we value lives in different legal contexts. What is one to say to that now? The achievement of “reflective equilibrium”—Rawls’s famous metaphorical description for what all forms of moral reasoning ultimately seek to accomplish, and taken by many to be as applicable to legal as to moral reasoning—is chiefly a matter of achieving consistency and casting out whatever is inconsistent. If the consistency constraint cannot be relied upon to do that, how is one to proceed?

### ***7. A Comparison with Robert Nozick’s Side-Constraint View of Rights***

Our model of the decision-maker maximizing subject to the need to rationalize his decision might remind some readers of Robert Nozick’s idea about rights as side constraints, and justly so. The two are intimately related and the relationship can shed quite a bit of light on the old questions raised by Nozick’s account (Nozick 1977).

Let us recall the details of Nozick’s view. He pointed out that rights violations can be incorporated into an ethical view in two fundamentally different ways. A consequentialist would do so by making them part of the overall good to be maximized. A deontological account, the approach favored by Nozick, would include them as side-constraints. The first approach might find it acceptable to

“violate someone’s rights when doing so minimizes the total amount of violation of rights in the society. For example, violating someone’s rights might deflect others from their intended action of gravely violating rights, or might remove their motive for doing so, or might divert their attention, and so on. A mob rampaging through a part of town killing and burning will violate the rights of those living there. Therefore, someone might try to justify his punishing another he knows to be innocent of a crime that enraged a mob, on the grounds that punishing this innocent person would help to avoid even greater violations of rights by others, and so would lead to a minimum weighted score for rights violations in the society.”

However,

“In contrast to incorporating rights into the end state to be achieved, one might place them as side constraints upon the actions to be done: don’t violate constraints C. The rights of others determine the constraints upon your actions. . . . This view differs from one that tries to build the side constraints C into the goal G. The side constraint view forbids you to violate these moral constraints in the pursuit of your goals; whereas the view whose objective is to minimize the violation of these rights allows you to violate the rights (the constraints) in order to lessen their total violation in the society.”

He acknowledges that there appears to be something quite strange about side constraints. “Isn’t it irrational to accept a side constraint C, rather than a view that directs minimizing the violations of C? ...If nonviolation of C is so important, should that be the goal? How can concern for the nonviolation of C lead to the refusal to violate C even when this would prevent other more extensive violations of C? What is the rationale for placing the nonviolation of rights as a side constraint upon action instead of including it solely as a goal of one’s actions?” His explanation is that “The moral constraints upon what we may do, I claim, reflect the fact of our separate existences. They reflect the fact that no moral

balancing act can take place among us; there is no moral outweighing of one of our lives by others so as to lead to a greater overall social good. There is no justified sacrifice of some of us for others.”

Now let’s turn to comparing our model with Nozick’s. To begin with, let’s clear away a distinction without a difference. Nozick permits maximization of preferences subject to not violating any rights constraints. Our model permits maximization subject to being able to rationalize the choice made. There is a prohibition in the first case and an affirmative requirement in the second. But of course each can really easily be recast in terms of the second. Taking unrationalizable, or unjustifiable, actions is simply another way of referring to actions that violate a constraint.

What is a crucial difference between our model and Nozick’s however, is the nature of those constraints. In our model there is no trace of deontological rights. There is not anything really that prohibits tradeoffs. The rationalizations can be eminently consequentialist, even utilitarian rules. (Subject, however, to one important caveat: The restriction cannot simply be that an action is permitted if and only if it maximizes overall utility or some other end-state consequence of an analytically similar character). In particular, it might actually permit an individual actor to take actions that would otherwise be prohibited but which minimize the overall number of constraints being violated. It would not for instance prohibit allowing someone to engage in an act that would not ordinarily be rationalizable by necessity if the overall consequence would be that enough others will be inhibited in their commission of nonrationalizable actions to generate a net gain in the number of such violations avoided.

Notice however, that our model has a paradoxical feature very much like Nozick’s. It is the feature we already drew attention to when exploring whether there was or was not anything paradoxical about the duress, necessity, self-defense and negligence cycles laid out above. Is it really rational to insist that that someone who is willing to accept smoke inhalation for the sake of a

manuscript not be required to do the same for the sake of not committing a serious crime, or to accept an equivalent risk for the sake of not attacking someone who has attacked him?

There is an interesting general lesson about rights to be found here. The objections usually made to rights-based arguments can be made almost as readily against any rule-based arguments. All rule-based systems, even if the rules are of a fairly consequentialist character, are open to the objections made to rights-based reasoning. That rather takes the wind out of criticisms of rights-based approaches. Unless one is willing to do away with all rule-based regimes, one is going to have to put up with such objections in one form or another. Now we don't want to overclaim here. There has after all been plenty of recognition that everything except outright act utilitarianism has a rights-based flavor to it. Still, what has been less recognized is that this will be true even if the rules in question are fairly consequentialist in character.

### ***8. Rights and Fairness in Sen, Kaplow and Shavell***

Sen's theorem established that there is a severe tension between granting people rights and abiding by the Pareto principle (Sen 1979, chp. 6). Doing both, the theorem showed, leads to a cycle. Kaplow and Shavell have argued for a generalization of Sen's theorem according to which it is not merely rights that stand in tension with the Pareto principle but all fairness based legal doctrines (Kaplow and Shavell 2002). In other words, the Pareto principle combined with fairness based doctrines leads to a cycle as well. Our argument here casts a curious and somewhat unexpected light on these two claims: it suggests that they have rather different implications than they have been generally thought to have (Sen 1979; Kaplow and Shavell 2002).

Now from one point of view our analysis simply further generalizes these results. It suggests that it is not merely rights, or fairness-based doctrines, but rules more generally that stand in tension with the Pareto principle. To be sure, we haven't talked about the Pareto principle so far, but simply

about a cycle that results from combining respect for rules and personal preferences. But of course it is no far distance to travel from this to a conflict between rules and the Pareto principle. All we need assume is that the maximization of the person's preferences helps him, and hurts no one else, at least so long as his preferences can be rationalized.

From another point of view, however, our analysis should lead one to wonder about the import of those results. What we have here demonstrated is that the law will be forced to live with cycles, no matter what, among other things, with cycles that involve the Pareto principle. In other words, trying to make legal doctrines less fairness- or rights-oriented will not eliminate, and perhaps not even much reduce, the existence of such cycles. In a way, then, the tension between rights, fairness and the Pareto principle may turn out to be inconsequential, at least in that special sense.

### ***9. Conclusion***

Choices that exhibit cycling are often taken to be obviously irrational. Yet, as we show here, agents who show even a minimal respect for rules are bound to engage in exactly those kinds of choices. Since rule-based reasoning is not generally regarded as irrational, neither, it seems, should cyclical choices. This produces a number of implications and questions. First, it casts doubt on many criticisms of legal doctrines and theories of law (such as Kaplow and Shavell's critique of fairness-based arguments, Sen's critique of the Pareto principle based on its incompatibility with rights, and critiques of rights theory based on the so-called deontological paradox) that argue that they are incoherent because they lead to inconsistent, that is, cyclical choices. Second, it suggests a way in which all legal rules are subject to circumvention to a much greater extent than generally appreciated. And third, it raises the very fundamental question as to when inconsistency is a valid basis for criticism and when it is not.

Appendix:

## 1. Mathematical Model of a Law-Abiding Decision Maker

Let  $A$  be a finite set of alternatives. A decision maker, called Dee, chooses from a set of  $B \subseteq A$  of feasible options. In addition to the feasibility constraint, Dee faces a rationalization constraint. Intuitively, an alternative  $x \in B$  can be rationalized if and only if Dee can justify this alternative to herself using some rationale. A rationale  $R^i$  is a way to compare alternatives. Formally, a rationale  $R^i$  is an asymmetric binary relation and  $x R^i y$  denotes that, by rationale  $R^i$ ,  $x$  is acceptable in the presence of  $y$ . Rationales can be, but need not be, transitive or complete and so, can accommodate reasons that relate to some, but not all, alternatives. Given an set of alternatives  $B$ , an option  $x \in B$  is rationalized by  $R^i$  when  $x R^i y$  for all  $y \in B$ ,  $y \neq x$ . That is, no physically feasible alternative  $y$  makes choosing  $x$  unacceptable according to rationale  $R^i$ . Let  $R = \{R^i, i=1, \dots, n\}$  be the set of all rationales available to Dee. An option  $x \in B$  is rationalizable in  $B$  if  $x$  can be rationalized by some rationale  $R^i \in R$ . Thus, an option can be rationalized is Dee can find some rationale that makes  $x$  an acceptable choice (even if according to all other rationales  $x$  is not rationalizable). Let  $L(B) \subseteq B$  be the set of rationalizable options in  $B$ . By definition, the complement of  $L(B)$  (in  $B$ ) are the options that cannot be rationalized.

This model can accommodate the idea that some options are rationalized because there is no compelling reason to forbid them. Formally, for such an option, say  $w$ , there is a rationale  $R^i \in R$  such that  $w R^i z$  for all  $z \in A$ ,  $z \neq w$ . We assume that the set of rationalizable options is always non-empty.

Dee has a preference order  $P$  (i.e., a transitive, asymmetric and binary relation) over all alternatives. So, Dee ranks all feasible alternatives and  $x P y$  denotes that Dee prefers  $x$  over  $y$ . Our main assumption is that Dee chooses the feasible alternative  $C(B) \in B$  that she prefers among all rationalizable alternatives. Formally,  $C(B) P y$  for all  $y \in L(B)$ . So, Dee solves a constrained optimization problem. That is, Dee optimizes  $P$  subject to  $L(B)$ .

Now consider the special case in which Dee is a law-abiding citizen. That is, Dee can rationalize any option that is legal and none that are illegal. Whether or not Dee fears punishment from breaking the law is not critical for this exercise. Dee is psychologically constrained to stay within the limits of the law. Thus, even if Dee is sure that she will not be discovered and/or punished by taking some illegal actions she still, for psychological and ethical reasons, can only rationalize legal options and so, chooses the legal option that she prefers. In this case,  $L(B)$  is the set of all legal options. This simplifies the analysis because the consequences of taking illegal acts need not be described. We now explore the logical structure of the legal constraint and its implications for decision making.

## 2. Cycles of a Law-Abiding Decision Maker

The formal model described above where Dee maximizes her preferences subject to legal constraints may seem logically analogous to standard economic theory where Dee maximizes her preferences subject to budget constraints. Indeed, in both cases Dee behaves as if she solves a constrained optimization problem of a preference order (or, equivalently, an utility function). As we mentioned, the fundamental difference lies in the logical structures of legal and budget constraints. Consider the following example. Assume that there are three options  $x$ ,  $y$  and  $z$  and the constraint  $L$  satisfies

$$x \in L(x,y) \text{ and } x \notin L(x,z)$$

This property (henceforth called pattern 1) is impossible if  $L$  is a budget constraint, but it is possible if  $L$  is a legal constraint. To see this, note that whether or not Dee can afford  $x$  depends only on  $x$ 's price and Dee's resources. So, the affordability of  $x$  is unaffected if good  $y$  is no longer available and is replaced with good  $z$  (although this change may affect the final purchase). In the case of legal constraints, pattern 1 is possible. This is a key reason why decisions in standard consumer theory are not cyclic, but decisions of law-abiding citizens may be.

The examples in Section 4 show why pattern 1 is possible in the case of legal constraints. Consider the example on duress. Assume that  $x$  is to significantly harm someone else,  $y$  is to endure severe pain, and  $z$  is to give up a valued manuscript. Now, we may have  $x \in L(x,y)$  because if the only feasible alternative to  $x$  is to endure severe pain then the duress defense may make  $x$  legally permissible. Formally, there is a rationale  $R^1 \in R$  such that  $x R^1 y$  and so,  $x \in L(x,y)$ . On the other hand, if  $z$  is now available then Dee may avoid doing harm to anyone by giving up a valued manuscript. So,  $x$  is no longer legal and so,  $x \notin L(x,z)$ .

Now consider the decision between enduring severe pain and giving up a valued manuscript. Given that there is no legal restriction forbidding pain endurance, it is legal to endure severe pain in order to keep the valued manuscript. That is, there is a rationale  $R^2 \in R$  such that  $y R^2 z$  and so,  $y \in L(y,z)$ . Hence, we have the following pattern (called pattern 2)

$$x \in L(x,y), y \in L(y,z) \text{ and } x \notin L(x,z).$$

Now assume that Dee ranks  $x$  above  $y$  and  $y$  above  $z$ . Given that the legal constraint  $L$  satisfies pattern 2, her optimal choices are cyclic, i.e.,  $C(x,y) = x$ ,  $C(y,z) = y$ ,  $C(x,z) = z$ .<sup>1</sup>

Dee's preferences are a strict order and she optimizes her preferences under legal constraints. Yet, her behavior is cyclic. Dee's cyclic behavior comes from pattern 2 in the legal constraint  $L$ . This pattern can be obtained when there are two rationales  $R^1 \in R$  and  $R^2 \in R$  such that  $x R^1 y$  and  $y R^2 z$ , but no rationale, say  $R^3$ , such that  $R^3 \in R$  and  $x R^3 z$ . The examples in section 4 on necessity, self-defense and negligence can be formalized in an analogous way to the example on duress.

### 3. Logic and Consistency in Law

---

<sup>1</sup> By assumption,  $L(x,z)$  is non-empty and so,  $x \notin L(x,z)$  implies that  $z \in L(x,z)$

Transitivity is often seen as one of the most basic tenet of rationality. Yet, it is, and must be, routinely violated by choices made in legal settings. As we pointed out, this may suggest a difficulty with legal arguments based on inconsistencies. However, these results do not show that the decision of a law abiding citizen do not follow any logical structure. They also do not imply that the law must necessarily be either inconsistent or lead to irrational choices. Rather, the argument suggests that what should be taken as evidence of irrationality depends on the underlying model of the behavior. Cyclic choice is irrational in standard economic theory because it is inconsistent with the idea that the preferred choice was always exercised. On the other hand, cyclic choices are not conclusive evidence of irrationality under rationalization theory because the theory can accommodate cycles. However, some choice patterns are ruled out by rationalization theory. Consider the following choices

$$C(x,y) = x ; C(x,y,z) = y ; C(x,y,z,w) = x.$$

These choices cannot be accommodated by rationalization theory. The choices  $C(x,y) = x$  and  $C(x,y,z) = y$  imply that Dee prefers  $x$  over  $y$ . Otherwise the choice of  $x$  over  $y$  is only possible if  $y$  cannot be rationalized in the presence of  $x$ . But  $C(x,y,z) = y$  implies that  $y$  can be rationalized in the presence of  $x$ . By analogous argument,  $C(x,y,z) = y ; C(x,y,z,w) = x$  imply that Dee must rank  $y$  above  $x$ . This contradiction shows that the choices above are inconsistent with rationalization theory.<sup>2</sup> Hence, the choice of a law abiding citizen are not arbitrary and do have an abstract logical structure. In particular, it must be that if  $C(x,y) = x$  and  $C(x,y,z,w) = x$  then  $C(x,y,z) \neq y$ .

---

<sup>2</sup> Rationalization theory cannot accommodate all choice patterns because there is a logical structure that the constraints  $L$  must satisfy. As shown in Cherepanov, Feddersen, and Sandroni, the following property must be satisfied by any constraints  $L$  in the rationalization model:

$$\text{If } B \subseteq B^* \text{ then } \{L(B^*) \cap B\} \subseteq L(B).$$

These choices are also inconsistent with the standard economy theory. So, while rationalization theory raises the bar for a demonstration of inconsistency, it does not raise the bar to the point that no evidence is sufficient to be considered inconsistent or irrational.

## References

American Law Institute. 1981. *Model Penal Code*. Philadelphia, Penn.: American Law Institute.

Berger, Paul D., and Gerald E. Smith. 1997. The Effect of Direct Mail Framing Strategies and Segmentation Variables on University Fundraising Performance. *Journal of Direct Marketing* 2: 30-43.

Chapman, Bruce. 2003. Rational Choice and Categorical Reasoning. *University of Pennsylvania Law Review* 151: 1169-1210.

Cherepanov, Vadim, Timothy Feddersen, and Alvaro Sandroni. 2008. Rationalization. Working Paper. Department of Economics, University of Pennsylvania, and Kellogg School of Management, Northwestern University.

Cooter, Robert. 1984. Prices and Sanctions. *Columbia Law Review* 84: 1523-1560.

Cooter, Robert, and Ariel Porat. 2001. Should Courts Deduct Normal Sanctions from Damages? *Journal of Legal Studies* 30: 401-422.

Kaplow, Louis, and Steven Shavell. 2002. *Fairness versus Welfare*. Cambridge, Mass.: Harvard University Press.

Kornhauser, Lewis, and Larry Sager. 1986. Unpacking the Court. *Yale Law Journal* 96: 82-118.

Nozick, Robert. 1977. *Anarchy, State, and Utopia*. New York: Basic Books.

Posner, Eric. 2002. *Law and Social Norms*. Cambridge, Mass.: Harvard University Press.

Sen, Amartya. 1979. *Collective Choice and Social Welfare*. San Francisco, Calif.: Holden-Day.

Snyder, Melvin L., Robert E. Kleck, Angelo Strenta, and Steven J. Mentzer. 1979. Avoidance of the Handicapped: An Attributional Ambiguity Analysis. *Journal of Personality and Social Psychology* 37: 2297-2306.

Spitzer, Matthew L. 1979. Multicriteria Choice Processes. *Yale Law Journal* 88: 717-780.

Stearns, Maxwell. 1977. *Public Choice and Public Law: Readings and Commentary*. Cincinnati: Aderson Publishing Company.