Inventive Retaliation:
Adam Smith, David Laitin, and the
Costs of Sustaining Social Norms

Barry R. Weingast*
Stanford University

Abstract

Maintaining social norms often requires that members of a community to undertake costly punishment. Under what circumstances is it rational for members to bear this cost? Laitin emphasizes some people invent clever retaliations so as to be admired for doing so (personal communications). The purpose of this note is to provide an answer by drawing on Adam Smith’s *The Theory of Moral Sentiments* (1759) to develop a framework in which clever retaliation is a rational response for some when others have violated social norms. Smith’s approach suggests a way of modeling the phenomenon identified by Laitin.

1. Introduction

Maintaining social norms often requires that members of a community to undertake costly punishment. Under what circumstances is it rational for members to bear this cost? Undertaking costly punishments is puzzling in that it is a public good from which all members of the community benefit. As is well-known, people have incentives to let others bear the costs of the public good; but when everyone does that, the community fails to provide the good.

David Laitin, one of the premier contemporary students of culture, emphasizes the importance for the maintenance of social norms that some people invent clever retaliations so as to be admired for doing so (personal communications). How should we model this “inventive retaliation,” especially since retaliation is costly and clever retaliation may involve higher costs than other retaliation?

* Senior Fellow, Hoover Institution, and Ward C. Krebs Family Professor Department of Political Science, Stanford University. The author gratefully acknowledges David Laitin and Glory Liu for helpful conversations.
The purpose of these notes is to provide an answer to this question. To do so, I draw on Adam Smith's *The Theory of Moral Sentiments* (1759) to develop a framework in which clever retaliation is a rational response for some when others have violated social norms. Smith's approach suggests a way of modeling the phenomenon identified by Laitin.

To summarize the idea, the argument has three components.

1. Smith’s theory of individual behavior in this book is based on self-interest tempered with the idea of *approbation*, namely, that people seek not only praise but to become known as being praiseworthy. [*TMS* cite]

2. Smith’s theory of justice is commutative, based on retaliation for bad offenses. [*TMS* cite]

3. Putting these two ideas together generates Laitin’s theory. Members of a community praise members who undertake costly punishment. Further, members of the community believe that individuals who regularly invent clever retaliations deemed appropriate for the circumstances are praiseworthy. When an inventive retaliator values this approbation more than the costs of punishment, punishment is credible.

Although I do not emphasize this point, this paper can be thought of as an application of behavioral economics. Ashraf, Camerer, and Lowenstein (2005) provide an elegant introduction to Adam Smith's ideas about behavioral economics.

These notes consider these three topics in turn.

**2. Approbation – and the Impartial Spectator**

Adam Smith argues that, to varying degrees, people all seek admiration and especially approbation from others. A major point is that this motive causes us to temper our natural self-centered nature, or in Smith's terms, our "self-love."²

The desire for approbation directly affects our behavior. Smith also emphasizes that this motive leads many to be benevolent – because they want to be known for being a benevolent person. As part of this approach, Smith develops the concept of the impartial spectator, the idea that to become praise-worthy, we must look at ourselves in the manner of an impartial observer. The impartial spectator helps people temper their self-centered nature, aiding their ability to see themselves – at least to a degree – as

---

² This section draws on the extensive literature on approbation and the impartial spectator, including: Macfie **, Raphael **
others see them. This allows people, at least to a degree, to adjust their behavior to gain approbation.

Smith opens *TMS* with this idea:

> How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it. [*TMS* 47]

More generally,

Man naturally desires, not only to be loved, but to be lovely; or to be that thing which is the natural and proper object of love. He naturally dreads, not only to be hated, but to be hateful; or to be that thing which is the natural and proper object of hatred. He desires, not only praise, but praise–worthiness; or to be that thing which, though it should be praised by nobody, is, however, the natural and proper object of praise. He dreads, not only blame, but blame–worthiness; or to be that thing which, though it should be blamed by nobody, is, however, the natural and proper object of blame. [*TMS* 208]

The love and admiration which we naturally conceive for those whose character and conduct we approve of, necessarily dispose us to desire to become ourselves the objects of the like agreeable sentiments, and to be as amiable and as admirable as those whom we love and admire the most. Emulation, the anxious desire that we ourselves should excel, is originally founded in our admiration of the excellence of others. Neither can we be satisfied with being merely admired for what other people are admired. We must at least believe ourselves to be admirable for what they are admirable. But, in order to attain this satisfaction, we must become the impartial spectators of our own character and conduct. We must endeavour to view them with the eyes of other people, or as other people are likely to view them. When seen in this light, if they appear to us as we wish, we are happy and contented. But it greatly confirms this happiness and contentment when we find that other people, viewing them with those very eyes with which we, in imagination only, were endeavouring to view them, see them precisely in the same light in which we ourselves had seen them. Their approbation necessarily confirms our own self–approbation. Their praise necessarily strengthens our own sense of our own praise–worthiness. In this case, so far is the love of praise–worthiness from being derived altogether from that of praise; that the love of praise seems, at least in a great measure, to be derived from that of praise–worthiness. [*208-09, emphasis added*]

In short, Smith argues that we temper our self-love because we want to be admired, and others rarely admire people who are wholly self-centered. To be the object of praise
– and, more importantly, being viewed as praise-worthy – requires active effort. The impartial spectator is a device useful for this task.

3. Smith’s Theory of Commutative Justice Based on Retaliation

An important element of Smith’s theory of justice is based on resentment.

Resentment seems to have been given us by nature for defence, and for defence only. It is the safeguard of justice and the security of innocence. It prompts us to beat off the mischief which is attempted to be done to us, and to retaliate that which is already done; that the offender may be made to repent of his injustice, and that others, through fear of the like punishment, may be terrified from being guilty of the like offence. [TMS 156]

Further,

the violation of justice is injury: it does real and positive hurt to some particular persons, from motives which are naturally disapproved of. It is, therefore, the proper object of resentment, and of punishment, which is the natural consequence of resentment. As mankind go along with, and approve of the violence employed to avenge the hurt which is done by injustice, so they much more go along with, and approve of, that which is employed to prevent and beat off the injury, and to restrain the offender from hurting his neighbours. The person himself who meditates an injustice is sensible of this, and feels that force may, with the utmost propriety, be made use of, both by the person whom he is about to injure, and by others, either to obstruct the execution of his crime, or to punish him when he has executed it. And upon this is founded that remarkable distinction between justice and all the other social virtues, which has of late been particularly insisted upon by an author of very great and original genius, that we feel ourselves to be under a stricter obligation to act according to justice, than agreeably to friendship, charity, or generosity. [TMS 157]

Fleischacker (2005, 151, emphasis added) explains these ideas: "This means both that a victim of harm will tend to resent and retaliate against harm to himself, and that 'every generous spectator' will not only approve of such conduct but will 'enter so far into his sentiments as often to be willing to assist him.' Even in a condition without civil government, says Smith, when 'one man attacks, or robs, or attempts to murder another, all the neighbours take the alarm, and I think that they do right when they run, either to revenge the person who has been injured, or to defend him who is in danger of being so' (TMS 158). '[R]etaliation seems to be the great law which is dictated to us by Nature,' he says (TMS 82), and it is nature which implant in us the impulse to resent injury, and to sympathize with the justified resentment of others. Legal systems merely formalize – and importantly, moderate – this natural impulse.”
Though the breach of justice, on the contrary, exposes to punishment, the observance of the rules of that virtue seems scarce to deserve any reward. ... Mere justice is, upon most occasions, but a negative virtue, and only hinders us from hurting our neighbour. The man who barely abstains from violating either the person, or the estate, or the reputation of his neighbours, has surely very little positive merit. **He fulfils, however, all the rules of what is peculiarly called justice, and does every thing which his equals can with propriety force him to do, or which they can punish him for not doing.** We may often fulfil all the rules of justice by sitting still and doing nothing. [TMS II.i.i.9, p 160, ea]

Smith next distinguishes between three different types of behavior:

[i] As every man doth, so shall it be done to him, and retaliation seems to be the great law which is dictated to us by Nature. Beneficence and generosity we think due to the generous and beneficent. ... [ii] The violator of the laws of justice ought to be made to feel himself that evil which he has done to another; and since no regard to the sufferings of his brethren is capable of restraining him, he ought to be over-awed by the fear of his own. [iii] The man who is barely innocent, who only observes the laws of justice with regard to others, and merely abstains from hurting his neighbours, can merit only that his neighbours in their turn should respect his innocence, and that the same laws should be religiously observed with regard to him. [TMS 160]

In short, justice requires retaliation. But how so when it is a costly public good?

4. Synthesis: The Theory of Inventive Retaliation

I begin with three premises. First, most social norms and most laws are enforced in a decentralized manner; state-coercion is only a small part of the enforcement of good behavior. Hadfield and I (2012) develop this point at length with respect to law. This premise is especially true for social norms whose violation carry no legal sanctions.

Second, decentralized, collective punishment is often costly. Smith's theory of approbation and the sources of benevolence fit here. Some people seek to be known as praiseworthy. They can enhance this reputation when they not only participate in retaliation but do so in clever and inventive ways.

Third, the world is ever-changing, with new circumstances arising (as North 2005 emphasizes). New circumstances often reveal ambiguities in the laws and rules in the sense that there are multiple ways to interpret the rule given the new circumstances. A dilemma arises because some interpretations define the behavior in the new circumstances as good, while others interpret it as bad. The dilemma implies that the normative implications of behavior in the new circumstance are ambiguous. Further, in
the presence of the new, ambiguous circumstances, opportunists may attempt to exploit the ambiguity, arguing the good interpretation.

As Hadfield and Weingast (2012) observe, one of the differences between norms and laws is the idea of stewardship. In rule of law countries, the courts, legislatures, and the bureaucracy all serve as stewards in the sense that they have the authority to make definitive statements about how to extend the existing rules into the new circumstances. This is a major feature of the common law, for example. Hadfield and Weingast model the problem of enforcement of rules as a coordination game. Legal rulings serve as focal points. Indeed, one of the chief tasks of the legal stewards is to extend the focal point in the face of ambiguities and new circumstances. In contrast, informal social norms typically lack official stewards and thus have a more difficult time adapting to new circumstances.

The lack of stewards for social norms raises a problem. How does a consensus arise about how to extend the focal point into the new circumstances when multiple interpretations arise? This is a major problem for decentralized norms.

Enter the clever, inventive retaliator. An inventive retaliator serves multiple purposes. First, the inventive retaliator helps provide the public good of punishment when punishment is costly. She does so in part – according to Smith's approach – because of the approbation she receives from others. She bears the cost of retaliation because doing so means she will receive approbation. Moreover, the more clever the retaliation, the greater the approbation.

Second, the inventive component of the inventive retaliator emerges as follows. Recall the problem of new and ambiguous circumstances where people rationally disagree about how to extend the existing logic into the new circumstances. In this case, the inventive retaliator serves to create (or extend) the focal point into the new circumstances in a decisive way. People admire this behavior not only because it may be clever and inventive, but because it solves a dilemma, one which may have been growing because people lacked the means of coordinating in the new circumstances.

Consider the "El Chamisal" dispute about the ambiguity of the border between the United States and Mexico (see Weingast 1995). It helps us understand the problem of new circumstances, but one lacking a clever inventor. Consider an Axelrod and Keohane (1985) world in which countries police each other through retaliation against violators of norms and treaties. With respect to Latin America, the only hope of policing the United States is for these countries to act as one, to consider a defection by the United States against one as a defection against all.

After the U.S. invasion of Mexico, the United States and Mexico set the border between them in the Hildago Treat of 1846 as the middle of the Rio Grande. In the early part of the 20th century, a section of the river moved south near El Paso. The United States immediate occupied the land, claiming the letter of the law: the "middle of the river" had
moved. Mexico argued the opposite: the border is where the river was at the time of the
treaty, not something impermanent. Hence an unresolved dispute arose. Was the
United States being opportunist? Hard to say, and rational people in different countries
could come to very different interpretations. The main implication is that, given the new
circumstances, Latin American countries were likely to fail to coordinate against the
United States.

This example illustrates the problems that arise for informal norms in the absence of a
steward. There exists no authoritative way for the Latin American community and the
United States to resolve this dispute when Latin American countries disagree about
whether to side with Mexico or the United States.

Ellickson (1991) gives some good examples of inventive behavior in the escalating
retaliation against the new cattle rancher who violates all the local norms and ignores all
the standard forms of local retaliation by neighbors.

5. Conclusions

Punishment for violations of social norms are often costly; when the costs are too high,
members of the community face a dilemma in that the norms are hard to sustain.

Laitin suggests that this problem is solved when a subset of individuals wish to become
known and praised for being an inventive retaliator, that is, they invent clever ways of
punishing members of the community who violate social norms. I applied Adam Smith’s
approach in *The Theory of Moral Sentiments* to explain why some people have an
incentive to undertake costly punishments. The more public is the punishment, the
greater the degree of approbation received by the inventive retaliator and the more
costly the punishment to the individual who violated social norms.

More generally, how Smith’s approach applies.
References


