University of Richmond Law Review

Volume 36 | Issue 2 Article 3

2002

Signaling or Reciprocating? A Response to Eric Posner's Law and Social Norms

Dan M. Kahan

SIGNALING OR RECIPROCATING? A RESPONSE TO ERIC POSNER'S *LAW AND SOCIAL NORMS*

*Dan M. Kahan**

There is an almost heretical disenchantment with law percolating within the legal academy. Conventional wisdom sees law as the natural solution to problems of collective action. When attaining some societal good—for example, a clean environment, a stock of useful technologies, a public education system, or a transportation infrastructure—depends on the willingness of individuals to behave in a manner that is not in their material interest, the law supplies incentives—such as tax abatements for nonpolluters, property rights for inventors, and punishments for tax cheats—that bring individual interests into alignment with collective ones. The problem, though, is that a regime of regulatory incentives is costly. Even when such regulations are sensible in content, the expense of administering them dissipates the social surplus generated by collective undertakings. And often they *aren't* sensible in content. Securing the enactment of regulatory law presents a collective action problem in its own right, one that relatively small, intensely interested groups are more likely to overcome than are members of the public generally. The result is the embarrassing spectacle of special interest politics. Law is thus a deeply imperfect, if not a hopelessly flawed, instrument of social governance.

This analysis supplies a large part of the impetus behind the emerging literature on *law and social norms*. This movement seeks to identify psychological and social dynamics that promote contributions to collective goods without the prodding—and hence without the pathologies—of regulatory incentives. Law might have a constructive role to play in fostering these behavioral

* Professor of Law, Yale Law School.

mechanisms, but otherwise it should simply get out of the way of their natural evolution.

The particular mechanism that animates Eric Posner's theory of law and social norms is *reputational signaling*.[1] Even when it is not in their immediate material interest to do so, many individuals, he argues, will decide to contribute to collective goods if and to the extent that doing so earns them a reputation for reliability that induces others to enter into profitable undertakings with them in the future. Ranging over topics as diverse as family law, shaming penalties, and flag desecration, Posner's is by far the most systematic and wide-ranging theory of the law and social norms to date.

But is it right? While admirable in its ambition and insightful in particular details, Posner's signaling model, I will argue, furnishes a poor vehicle for the social norms and law project.

My analysis will focus on three criteria that I believe are appropriately used to judge any theory of law and social norms. The first is *behavioral realism*: for a theory to be reliable, it must rest on behavioral dynamics that we have good reason—from empirical testing or otherwise—to believe are grounded in fact. The second criteria is *political feasibility*: even when based on empirically supportable premises, a social norms theory, to be useful, must generate determinate and politically attainable policy prescriptions. The final criteria is *moral acceptability*, which refers to the compatibility of the theory and its prescriptions with our basic normative commitments, particularly the ones that we hold as citizens of a liberal democracy. Posner's signaling model fairs poorly under all three of these standards.

Because you cannot beat somebody with nobody, I will combine my critique of Posner's theory with a limited defense of an alternative one based on the dynamic of *reciprocity*. The reciprocity theory holds that individuals in collective action settings behave not like rational wealth maximizers but rather like moral and emotional reciprocators.[2] When they perceive that other individuals are voluntarily contributing to public goods, most individuals are moved by honor, generosity, and like dispositions to do the

1. ERIC A. POSNER, LAW AND SOCIAL NORMS 18–21 (2000).

2. For further discussion regarding the reciprocity theory, see Dan M. Kahan, *Trust, Collective Action, and Law*, 81 B.U. L. REV. 333, 335–39 (2001).

same. When, in contrast, they perceive that others are shirking or otherwise taking advantage of them, individuals are moved by resentment and pride to withhold their own contributions and even to retaliate if possible. The reciprocity theory implies that because individuals behave in this fashion, the most effective means to promote cooperative behavior in collective action settings is to promote *trust*—the shared belief that others can in fact be counted on to contribute their fair share to public goods, whether or not doing so is in their material self-interest. Indeed, the reciprocity theory warns that incentives can often backfire by implying that most individuals are *not* inclined to contribute to collective goods voluntarily. The reciprocity theory, I will suggest, offers greater behavioral realism, greater political feasibility, and greater moral acceptability than does Posner's signaling model.

I. BEHAVIORAL REALISM

As with nearly any social norms theory, the best way to understand Posner's is in opposition to the conventional theory of collective action. That theory, most famously associated with Mancur Olson,[3] assumes that individuals behave like rational wealth maximizers. Such individuals will rarely find it worth their while to contribute resources to attaining collective goods—those the benefits of which are, absent exclusionary devices, open to enjoyment by all—but instead will free-ride on the contributions that others make to them. As a result, too few individuals will contribute sufficiently, and the well-being of the group (including society at large) will be frustrated. The only way to overcome this dilemma is to furnish external incentives—either subsidies or penalties—that bring the interests of individuals into alignment with those of the groups to which they belong. From criminal law to environmental law, from tax fraud to business fraud, from regulation of the professions to regulation of the Internet, this is the story that animates American policymaking.

The foundation of Posner's challenge to the conventional view is that theory's failure to assess how *repeat interactions* affect behavior within collective action settings.[4] In a single encounter—a

3. *See generally* MANCUR OLSON, THE LOGIC OF COLLECTIVE ACTION (1965).
4. POSNER, *supra* note 1, at 15–18.

"one-shot prisoner's dilemma"—rational wealth maximizers might experience the collectively self-defeating incentives posited by Olson and other conventional theorists. But things will likely be different where those individuals anticipate encountering one another under such circumstances over and over again. In that case, each individual might perceive that his or her decision to contribute or free-ride in one meeting could influence the likelihood that the other individual will do the same in their next meeting. If so, they might both choose to forego the short-term material advantage of free-riding in the present in the expectation of reaping the long-term material benefits associated with mutual cooperation across the run of future interactions. This insight was systematically developed by the political scientist Robert Axelrod. Using computer simulations and other ingenious proofs, Axelrod established that even rational wealth maximizers will converge on a stable pattern of cooperative behavior if they expect an indefinitely long sequence of future interactions with individuals whose behavior in collective action settings they are in a position to monitor and keep track of.[5]

Posner's distinctive refinement of Axelrod's theory is his account of *reputational signaling*. Posner assumes that individuals will find other individuals desirable as partners in profitable collective undertakings to the extent that they perceive them to have *low* rather than *high discount rates*.[6] Individuals who value future payoffs as much or nearly as much as present ones are more likely than those who excessively discount the future to forego the immediate advantages of shirking or cheating. They are more likely to do this in the interest of reaping the long-term benefits associated with access to future trading opportunities. Accordingly, individuals will seek to acquire a reputation for subordinating short-term to long-term payoffs by engaging in behavior that convincingly demonstrates such a disposition in the eyes of others. According to Posner, a social norm is just a widespread pattern of behavior that signals a low rather than a high discount rate in this manner.[7] If contributing to collective goods of various sorts (among other things) becomes such a signal, then a robust regime of social norms might minimize the necessity of relying on

5. *See* ROBERT M. AXELROD, THE EVOLUTION OF COOPERATION (1984).
6. *See* POSNER, *supra* note 1, at 19–21.
7. *Id.* at 22–23.

law—with all its attendant costs and pathologies—as a device for solving problems of collective action.[8]

That's Posner's story. The question I now want to consider is whether it is a *behaviorally realistic* or factually reliable one. Posner does not present social science data to support his reputational signaling model of social norms; he candidly acknowledges, in fact, that no such data exist. Nevertheless, he invites us to credit his account, at least provisionally, on the basis of its power to furnish plausible explanations for a wide range of readily observable phenomena and to generate empirically testable hypotheses.[9]

I do not object in principle to the form of pragmatic conjecture that characterizes Posner's argument. Indeed, I embrace it as frequently the most politically useful and morally responsible mode of analysis for guiding public policy choices, which we are often constrained to make on the basis of our best current understanding of human behavior, whether or not that understanding has been subjected to the rigors of social science testing.[10] Nevertheless, I believe Posner's signaling model, as it stands, *fails* this pragmatic test and should *not* be trusted to guide public policy even provisionally for four distinct reasons.

First, whatever informal or casual explanatory power Posner's account appears to have is attributable to its *overdetermination*. A theory can be said to be overdetermined when it furnishes a menu of opposing behavioral mechanisms that are sufficiently abundant to account for essentially any phenomena as well as its negation. In that circumstance, the theory does not generate "explanations" at all; it merely supplies a convenient set of storytelling templates that allow the theorist to rationalize *ex post* whatever existing facts she encounters and to justify whatever policy prescriptions she chooses *ex ante*. Rational choice theories are notoriously vulnerable to this defect.[11]

8. *Id.*

9. *Id.* at 36–38.

10. *See generally* Tracey L. Meares & Dan M. Kahan, *Law and (Norms of) Order in the Inner City*, 32 LAW & SOC'Y REV. 805 (1998).

11. *See generally* JON ELSTER, NUTS AND BOLTS FOR THE SOCIAL SCIENCES 7-10 (1989); Jon Elster, *A Plea for Mechanisms*, *in* SOCIAL MECHANISMS: AN ANALYTICAL APPROACH TO SOCIAL THEORY 45 (Peter Hedström & Richard Swedberg eds., 1998); Jeanne L. Schroeder, *Just So Stories: Posnerian Methodology*, 22 CARDOZO L. REV. 351

Posner's is no exception. The signaling model explains why *costly* patterns of behavior, "like gift-giving, consumption of expensive goods, and shunning of people with certain characteristics," become norms: precisely because such behaviors involve sacrifice of material wealth or opportunities to accumulate it, they convey that those who engage in them "value future payoffs more than bad types [i.e., those with high discount rates] do."[12] But some norms, "like combing [one's] hair in one way rather than another," impose only trivial costs.[13] No problem, the signaling model explains conformity to them as well: "people engage even in *cheap* actions . . . because their deviation from the norm will be punished by *others* who seek to signal their types by taking the costly action of shunning people who act in an unusual way."[14]

The signaling model readily explains why individuals slavishly *conform* to inconvenient but widespread patterns of behavior. If they did not, members of the social mainstream would infer that they lack discipline or willpower—in other words, that they have high discount rates—and would therefore exclude them from profitable collective undertakings.

Yet the signaling model can also explain why others conspicuously *flout* mainstream norms. The wealthy and powerful are not dependent on the collaboration of others and thus frequently feel no compunction against "violat[ing] common norms against ostentation, waste, and civility"; and "because people sometimes infer that those who violate norms must be wealthy and powerful, there is an independent incentive for people to violate norms—as a way of signaling their wealth and power," attributes that make them attractive trading partners after all.[15] Members of deviant subgroups will also ostentatiously defy mainstream norms—possibly by engaging in behavior that might connote excessive present-mindedness (i.e., a high discount rate), such as body-piercing, law breaking, or drug use—for signaling purposes.[16] Precisely because such behavior predictably *alienates* them from

(2001).

12. POSNER, *supra* note 1, at 25.
13. *Id.*
14. *Id.* at 25–26.
15. *Id.* at 28.
16. *Id.* at 28–29.

the mainstream, those who engage in it demonstrate the depths of their commitment to the deviant subgroup and thus make themselves appear reliable to other deviants, thereby maximizing "future payoffs from people outside mainstream society."[17]

In short, "anything"—costly sacrifices or trivial gestures, conformity or rebellion—"can be a signal."[18] But if *anything* can be a signal, then the signaling model can be used to tell whatever story the modeler wants.

Notwithstanding its malleability, Posner's attempt to reduce all norms to signaling behavior generates at least one consistent result that seems completely at odds with "common intuitions and data"[19]—namely, the correlation of zealotry with materialist values. This is the second reason to doubt the theory's behavioral realism.

Social norms notoriously underwrite all manner of intolerance and persecution. For Posner, this aspect of social norms, too, derives from the power of conformity to signal one's fitness as a potential trading partner to members of the social mainstream. By denouncing flag burners and other political subversives,[20] discriminating against racial minorities,[21] and persecuting homosexuals,[22] individuals conspicuously forego immediate opportunities to profit from interactions with them, and thus signal a low discount rate that redounds to their long-term material benefit.[23]

This story, which figures throughout *Law and Social Norms*, is just plain weird. Casual observation furnishes little reason to believe that those who submit to virulent frenzies of intolerance end up with the greatest business opportunities. On the contrary, someone who shows up, say, for a job interview or for a meeting with a prospective creditor espousing "white pride," homophobia, or even fervent patriotism is much more likely to alienate than to

17. *Id.* at 29.

18. *Id.* at 25.

19. *Id.* at 36.

20. *Id.* at 115–17.

21. *Id.* at 133–40.

22. *Id.* at 84–86, 127.

23. Of course, others choose to engage in flag burning to signal that they "reject[] the values of the majority," a costly gesture that "reduces one's opportunities to deal with those who respect the flag" and thus "serve[s] to reveal that one has a low discount rate" and is therefore a suitable trading partner for members of the subversive subcommunity. *Id.* at 117. Again, "anything can be a signal." *Id.* at 25.

impress a potential trading partner—even one who (more quietly) shares his particular chauvinism. A more sensible strategy for marketing oneself in the world of commerce is to show that one loves money so much that one is willing to do business with all comers; think of Michael Eisner's sponsorship of "Gay Pride Day" at Disney World over the howls of protest and dire warnings of the religious right.[24] Indeed, capitalism is famously (if unevenly) praised for its tendency to supplant group identification with more individualistic, bourgeois tastes and thus to quiet historic sources of social factionalization and strife.[25] A behavioral theory that seems to get this feature of our social experience so decidedly wrong—as Posner's does—hardly seems reliable.

The third reason to doubt the behavioral realism of Posner's theory has to do with the coherence of his central claim that a person's conformity to norms signals her "discount rate." According to a vast body of empirical literature, the phenomenon of a generic "discount rate" just does not exist.[26] The concept of a "discount rate" characterizes an individual's preference for immediate over future wealth. But since 1937, when Paul Samuelson first advanced the idea that individuals have such a preference,[27] sustained empirical research by economists and others have failed to demonstrate behavior consistent with it. Instead of exhibiting any generic discount rate, individuals' preferences for immediate over future wealth have been shown to be highly context specific. Indeed, a plausible interpretation of the empirical literature is that the degree to which individuals are willing to sacrifice present gains for future ones is governed by social norms that define when the seeking of immediate gratification is morally acceptable.[28] If so, discount rates cannot be said to cause social norms, as Posner posits; rather, norms cause (diverse, context-specific)

---

24. *See* Thomas B. Edsall, *Forecasting Havoc for Orlando: On TV, Robertson Says Display of Gays' Flags Invites Disaster*, WASH. POST, June 10, 1998, at A11.

25. *See* ALBERT O. HIRSCHMAN, THE PASSIONS AND THE INTERESTS: POLITICAL ARGUMENTS FOR CAPITALISM BEFORE ITS TRIUMPH (1977); STEPHEN HOLMES, PASSIONS AND CONSTRAINT: ON THE THEORY OF LIBERAL DEMOCRACY (1995).

26. *See, e.g.*, Richard H. Thaler & George Loewenstein, *Intertemporal Choice*, *in* RICHARD H. THALER, THE WINNER'S CURSE: PARADOXES AND ANOMALIES OF ECONOMIC LIFE 92, 92–106 (1992).

27. *See generally* Paul A. Samuelson, *Some Aspects of the Pure Theory of Capital*, 51 Q.J. ECON. 469 (1937).

28. *See generally* Shane Frederick, George Loewenstein & Ted O'Donoghue, *Time Discounting: A Critical Review*, J. ECON. LITERATURE (forthcoming 2002).

discount rates. In any event, the impossibility of specifying any individual's generic preference for immediate over future wealth makes it impossible to accept Posner's invitation to test his theory with data that correlates "conformity to social norms" with the "independent variable" of "discount rate."[29]

Fourth and finally, Posner's model is inconsistent with a wealth of empirical data on how individuals actually *do* behave in collective action settings. Economists and social psychologists have developed a diverse array of constructs—the "investment game,"[30] the "ultimatum game,"[31] the "moonlighting game,"[32] and others[33]—to simulate collective action problems in the laboratory. They have also extensively tested and measured collective action dynamics in field experiments and various real world environments.[34] Such studies consistently show that most individuals in collective action settings tend to adopt a conditionally cooperative stance, contributing to collective goods if and to the extent that they perceive the others are inclined to do the same.[35] This finding contravenes the conventional theory, which predicts that individuals, as rational wealth maximizers, will withhold their contributions no matter what they expect others to do, generating an equilibrium of universal free-riding.

But this same literature also contravenes the logic of Posner's reputational signaling model. In the laboratory and other settings, individuals will behave like conditional cooperators even when they play one-shot or finitely repeated games under conditions of anonymity. Designed to exclude an Axelrodian explanation, these conditions also rule out a Posnerian one for the willingness of individuals to engage in conditional cooperation in

---

29. POSNER, *supra* note 1, at 36.

30. *See* Ernst Fehr & Simon Gächter, *Fairness and Retaliation: The Economics of Reciprocity*, 14 J. ECON. PERSP. 159, 162 (2000).

31. *See generally* Colin Camerer & Richard H. Thaler, *Ultimatums, Dictators and Manners*, 9 J. ECON. PERSP. 209 (1995).

32. *See generally* Klaus Abbink, Bernd Irlenbusch & Elke Renner, *The Moonlighting Game: An Experimental Study on Reciprocity and Retribution;* 42 J. ECON. BEHAV. & ORG. 265 (2000).

33. *See* Armin Falk & Urs Fischbacher, A Theory of Reciprocity 30 (Feb. 1999) (unpublished manuscript, on file with author).

34. *See generally* Elinor Ostrom, *Collective Action and the Evolution of Social Norms*, 14 J. ECON. PERSP. 137 (2000).

35. *See generally* Dan M. Kahan, The Logic of Reciprocity: Two Theories of Collective Action (2002) (unpublished manuscript, on file with author).

actual collective action settings. For where they do not anticipate dealing with each other repeatedly and indefinitely, and where they in fact cannot keep track of each other's past behavior, individuals who behave like simple rational wealth maximizers have no incentive—indeed, no capacity—to reward cooperation with cooperation (or defection with defection) for the sake of developing a reputation.

The findings of the empirical literature make sense only if we assume that individuals in collective action settings behave like moral and emotional reciprocators rather than mere wealth maximizers. Moral and emotional reciprocators behave like conditional cooperators *not* because they believe that doing so will enrich them materially but because they value the experience of behaving cooperatively for its own sake, so long as they are not being exploited by others when they do so. Consistent with ordinary experience, susceptible of empirical validation, and in fact amply supported by existing social science data, this reciprocity dynamic clearly dominates Posner's reputational signaling theory along the dimension of behavioral realism.

II. POLITICAL FEASIBILITY

To be useful, a social norms theory must not only be behaviorally realistic but also capable of generating determinate and sensible policy prescriptions. How does Posner's reputational signaling model rate on this scale?

A good test is the problem of *tax evasion*. Tax compliance is the archetypical collective action problem for public policy. Society collects taxes to finance a variety of goods—from education to highways to national defense—that benefit its members collectively but that self-interested individuals lack sufficient incentive to produce through voluntary transactions. Yet if we assume that individuals behave like rational wealth maximizers, then we should expect them to refrain from voluntarily paying their taxes as well. Accordingly, the conventional theory of collective action prescribes the use of material incentives—in the form of penalties for tax evasion—to induce compliance.[36]

36. *See generally* Michael G. Allingham & Agnar Sandmo, *Income Tax Evasion: A Theoretical Analysis*, 1 J. PUB. ECON. 323 (1972).

The assumption that individuals decide to comply or evade paying their taxes on the basis of a simple cost-benefit analysis, however, is embarrassingly ill-supported by empirical evidence. Tax compliance rates—which vary dramatically across nations—seem to bear no connection to enforcement levels.[37] For example, tax cheats face a much higher expected penalty in many European nations than they do in the United States, yet the United States enjoys a higher compliance rate.[38]

Tax scholars attribute such cross-national variance to differences in social norms, or "tax climates," across nations.[39] Whereas evasion is viewed as a breach of an important civic obligation in the United States, it is viewed as a fairly trivial regulatory offense—akin to a traffic violation—in many European ones.[40] These differences in norms have been confirmed by laboratory experiments, which show that Americans are significantly *less* likely to evade taxes whether the expected penalty for evasion is high or low than are, say, Spaniards, who freely underreport their taxes whenever the expected return for doing so is positive.[41]

Indeed, far from promoting compliance, simply increasing the penalties for evasion has been shown to undermine it, at least in societies that otherwise enjoy relatively compliant norms. One study, for example, found that taxpayers who were exposed to information emphasizing the severity of tax-evasion penalties claimed *more* deductions than did similarly situated taxpayers exposed either to a moral appeal or to no information at all.[42] Another study found that individuals who were shown actual press accounts of an IRS plan to attack the "tax gap" with stepped-up auditing displayed a *weaker* commitment to paying their own taxes.[43]

---

37. Kahan, *supra* note 2, at 341.

38. *Id.*

39. *See* FRANK A. COWELL, CHEATING THE GOVERNMENT: THE ECONOMICS OF EVASION 102–03 (1990).

40. Kahan, *supra* note 2, at 341.

41. *See* James Alm, Isabel Sanchez & Ana de Juan, *Economic and Noneconomic Factors in Tax Compliance*, 48 KYKLOS 3, 13–14 (1995).

42. *See* Richard D. Schwartz & Sonya Orleans, *On Legal Sanctions*, 34 U. CHI. L. REV. 274, 298–99 (1967).

43. *See* Steven M. Sheffrin & Robert K. Triest, *Can Brute Deterrence Backfire? Perceptions and Attitudes in Taxpayer Compliance*, *in* WHY PEOPLE PAY TAXES: TAX COMPLIANCE AND ENFORCEMENT 193, 212–13 (Joel Slemrod ed., 1992).

Not surprisingly, Posner sees the signature of reputational signaling in the contribution that norms make to tax compliance. Scrupulously honoring one's tax obligations, Posner argues, is another way to demonstrate that one has a "low discount rate" and is thus a suitable trading partner.[44] Indeed, according to Posner, the value of tax compliance as a signal of this sort is greatest when the penalty is set well *below* the expected gain from evasion, for that's when it is most obvious that a compliant taxpayer is foregoing the temptation to cheat for immediate profit. As a result, a nation, such as the United States, can achieve high rates of compliance, motivated by signaling, notwithstanding low penalties. Nations that foolishly raise their penalties eviscerate the signaling value of compliance and can thus end up with *more* evasion on net.[45]

Although Posner's story "fits" the tax compliance data, (remember: "anything can be a signal"[46]), it furnishes a singularly unconvincing explanation for it. In order for any species of behavior to signal something about a person's discount rate, that behavior must be readily *observable*. In the United States, at least, whether a person is complying with her tax obligations *can't* be observed by members of the public generally. There is no public registry of individual tax returns, which are in fact protected from disclosure by privacy legislation.[47]

Indeed, a taxpayer deemed to have underpaid her tax by an IRS auditor can avoid *any* public record of her deficiency simply by settling with the government before the commencement of formal administrative or legal proceedings.[48] Thus, far from facilitating the positive signal of a clean tax record, United States law deliberately ambiguates it.

Of course, individuals are perfectly free to make their tax returns available for inspection, either by members of the public generally or by potential trading partners in particular. That's

---

44. *See* Eric A. Posner, *Law and Social Norms: The Case of Tax Compliance*, 86 VA. L. REV. 1781, 1789–90 (2000).

45. *See id.* at 1790–91.

46. POSNER, *supra* note 1, at 25.

47. *See* 26 U.S.C. § 6103(a) (1994).

48. Section 6103 of the Tax Code has been construed to cover information generated in IRS investigations, including audits. *See, e.g.*, Mallas v. United States, 993 F.2d 1111, 1122–24 (4th Cir. 1993).

presumably exactly what they would do if compliance were regarded as a valuable signal of their fitness for trade. But there is in fact no such norm of spontaneous tax disclosure in our society. Indeed, anyone who showed up at a commercial negotiation eager to display his or her latest tax returns would probably be regarded not as a trustworthy business partner but as some kind of freak.

Again, reciprocity furnishes a more plausible account of the contribution that norms make to tax compliance. Experimental studies show that whether an individual is likely to evade her tax obligations is strongly predicted by whether those around her express an intention to comply or not.[49] Likewise, the higher an individual believes the rate of tax-cheating to be in general, the more likely he or she is to cheat too.[50] These findings are exactly what one would expect were tax compliance and evasion motivated by reciprocal dispositions. Moral and emotional reciprocators take pride in contributing their fair share to public goods, but deeply resent being taken advantage of. With tax collection, as with other collective action settings, the extent to which others appear to be contributing to the good in question determines which of these sensibilities—pride or resentment—comes into play. If most other individuals seem to be paying their taxes, then the decision to pay hers gratifies a reciprocator's desire to be a good citizen. In contrast, if most individuals appear to be evading paying taxes, then complying will make her feel like a sucker.

The phenomenon of national "tax climates" also fits the reciprocity theory. Because individuals are reciprocators, their decisions in a collective-action setting feed on and reinforce each other, generating multiple high- and low-cooperation equilibria independent of the payoffs associated with cooperating or defecting. Accordingly, we should not expect, under the reciprocity theory, to see any fixed correspondence across nations between the

---

49. *See, e.g.*, Marco R. Steenbergen, Kathleen M. McGraw, & John T. Scholz, *Taxpayer Adaptation to the 1986 Tax Reform Act: Do New Tax Laws Affect the Way Taxpayers Think About Taxes?*, *in* WHY PEOPLE PAY TAXES: TAX COMPLIANCE AND ENFORCEMENT 9 (Joel Slemrod ed., 1992).

50. *See* Robert B. Cialdini, *Social Motivations to Comply: Norms, Values, and Principles*, *in* 2 TAXPAYER COMPLIANCE 200, 215 (Jeffrey A. Roth & John T. Scholz eds., 1989); James P. F. Gordon, *Individual Morality and Reputations Costs as Deterrents to Tax Evasion*, 33 EUR. ECON. REV. 797, 804 (1989); Steven Klepper & Daniel Nagin, *The Criminal Deterrence Literature: Implications for Research on Taxpayer Compliance*, *in* 2 TAXPAYER COMPLIANCE, *supra*, at 126, 144; Sheffrin & Triest, *supra* note 43, at 214.

expected penalty for evasion and compliance levels. If individuals believe that those around them are inclined to pay their taxes, they will (as a result of guilt, shame, pride and the like) be more likely to comply, thereby strengthening the perception that individuals are generally inclined to pay. If, in contrast, individuals believe that those around them are inclined to evade, resentment will inhibit them from complying, thus strengthening the perception that most individuals are inclined to cheat. In other words, what we should expect to see under the reciprocity theory is exactly what we do see—viz., competing and relatively durable norms toward tax compliance.[51]

The reciprocity theory not only furnishes a convincing explanation for the phenomenon of tax evasion; it also suggests a novel theory for combating it: the promotion of trust. If the government can make individuals believe that others are inclined to comply with their obligations, then individual taxpayers—as reciprocators—should be more inclined to pay themselves. This insight was borne out by a study sponsored by the Minnesota Department of Revenue.[52] In that study, researchers sent letters to a group of individuals stating that tax compliance rates were in fact much higher than what public opinion polls suggested citizens believed them to be.[53] Those individuals thereafter reported more income and claimed fewer deductions than did individuals in a control group.[54] Consistent with the reciprocity theory of collective action—and at odds with the conventional rational choice one—the Minnesota study also found that simply advising taxpayers that others were inclined to comply was more cost-effective than the threat of an audit![55]

Indeed, the reciprocity theory helps to explain why such threats have sometimes been shown to backfire. When the IRS engages in dramatic gestures to make individuals aware that it is

51. *See* Sheffrin & Triest, *supra* note 43, at 194–95 (suggesting interdependence of taxpayer decisionmaking should generate multiple behavioral equilibria); *see also* COWELL, *supra* note 39, at 112–13 (developing theoretical model predicting multiple compliance equilibria based on interdependence of taxpayers' decisions to evade).

52. *See* STEPHEN COLEMAN, MINNESOTA DEPARTMENT OF REVENUE, THE MINNESOTA INCOME TAX COMPLIANCE EXPERIMENT: STATE TAX RESULTS 5–6 (Apr. 1996), *available at* http://www.taxes.state.mn.us/reports/complnce.pdf.

53. *See id.*

54. *See id.* at 18–19.

55. *See id.* at 24–25.

redoubling its efforts to catch and punish tax evaders, it also causes individuals to infer that more taxpayers than they thought are choosing to cheat.[56] This inference in turn triggers a reciprocal motive to evade, which dominates the greater material incentive to comply associated with the higher than expected penalty. Because it misunderstands the contribution that social norms make to tax evasion, the conventional strategy suggests a self-defeating strategy for dealing with it.

But because it misunderstands the *reciprocal nature* of these norms, Posner's theory suggests a prescription that is similarly ill-conceived. As I have argued, the unobservability of tax compliance makes reputational signaling an unconvincing explanation for widespread compliance in the United States. But if one believed that individuals *could* be motivated to comply to send a signal of their low discount rates, and hence of their desirability as trading partners, then the right policy would presumably be to make noncompliance as transparent as possible. It would, in that case, make sense to repeal the provisions that now permit taxpayers to shield adverse audit findings from public view. Indeed, to promote signaling, it would probably be a good idea to institute procedures for publicizing the identities of delinquents as widely and dramatically as possible, possibly through shaming penalties.[57]

The logic of reciprocity, however, suggests that such reforms would likely be a disaster. Campaigns to broadcast the names of tax evasion cheats, like high-profile enforcement crack downs, could well induce individuals to revise upward their estimation of the extent of tax evasion. If so, such publicity would predictably generate resentment and a *negative* reciprocal motivation to evade in kind. The result would be not more tax compliance but less.

The reputational signaling model generates a politically naïve prescription about tax evasion. And there is little reason to think that it would do any better elsewhere.

56. *See id.* at 25.

57. *See generally* Dan M. Kahan & Eric A. Posner, *Shaming White-Collar Criminals: A Proposal for Reform of the Federal Sentencing Guidelines*, 42 J.L. & ECON. 365 (1999).

III. MORAL DESIRABILITY

The last criterion appropriate for judging social norm theories such as Posner's is *moral desirability*. Even assuming they are behaviorally realistic and generate worthwhile prescriptions, are the programs they underwrite consistent with our basic moral commitments?

To see what's at stake here, it helps to reformulate the impetus for the social norms project in slightly broader terms. What connects all of the social norm approaches is their responsiveness to what might be called *the paradox of liberalism*. The genius of liberal institutions is their power to detach us from divisive group identities and commitments. When we are conditioned to think of ourselves as wealth seekers, we stop thinking of ourselves as members of particular classes, as descendants of particular races or ethnicities, and as adherents to particular religious faiths. Accordingly, we stop thinking of those of *other* classes, races, ethnicities, and faiths as the bearers of status claims that compete with our own, and start to think of them instead as the bearers of *preferences* and *interests* that create opportunities for profitable trades and wealth-maximizing political arrangements.[58]

But if we do indeed think of ourselves as *mere* wealth seekers, it is unclear that we will succeed in exploiting these opportunities. As the conventional theory of collective action tells us, individuals who relentlessly pursue their material self-interest will behave in ways that systematically frustrate their collective well-being, both in markets and in politics. Using law to compel individuals to behave cooperatively in these settings is not a fully satisfactory solution, for the costly maintenance of such a regime dissipates the wealth associated with free trade, not to mention the liberty secured by liberal political institutions. Precisely because it disconnects us from our commitment-laden group identities, liberalism seems to leave us without the minimal moral capacity to commit to the economic and political institutions that a liberal society presupposes.

The social norm project attempts to remedy this defect in liberalism. Its aim is to identify psychological and social mechanisms

58. *See* HIRSHMAN, *supra* note 25, at 4.

that dispose individuals to behave consistently with their collective interests without legal prodding.

The problem, though, for most of the social norm programs is that they seem to assume the vitality of the group commitments that liberalism displaces. This is most obviously true for "communitarian" theories like Michael Sandel's[59] and even Robert Putnam's,[60] which advocate cultivation of the partisan attachments of ethnicity, faith, and locale. It is also true for proposals like Richard McAdams's,[61] as well as my own,[62] to harness "shame" and "stigma," which regulate individuals by threatening to reassign them from a respected high-status class to a despised low-status one.

But it is even true, ironically, for seemingly hard-edged rational choice theories like Robert Ellickson's, Lisa Bernstein's, and Eric Posner's that emphasize mechanisms like "tit-for-tat,"[63] "reputation,"[64] and "signaling."[65] Descended from Axelrod's brilliant revision of the conventional theory, all of these approaches depend on very specialized conditions: rational wealth maximizers will have an incentive to reciprocate cooperation with cooperation and free-riding with costly forms of punishment *only* if they anticipate dealing indefinitely with particular individuals whose behavior they can observe and keep track of. Those fairly intimate conditions are not present in the tax-compliance setting, for example; nor are they present in most of the other *n*-person prisoner's dilemmas that are the subject of public policymaking in

59. *See* MICHAEL J. SANDEL, DEMOCRACY'S DISCONTENT: AMERICA IN SEARCH OF A PUBLIC PHILOSOPHY 231–32 (1996).

60. *See* ROBERT D. PUTNAM, BOWLING ALONE: THE COLLAPSE AND REVIVAL OF AMERICAN COMMUNITY (2000).

61. *See* Richard H. McAdams, *The Origin, Development, and Regulation of Norms*, 96 MICH. L. REV. 338, 365–66 (1997).

62. *See generally* Dan M. Kahan, *What Do Alternative Sanctions Mean?*, 63 U. CHI. L. REV. 591 (1996).

63. *See* ROBERT C. ELLICKSON, ORDER WITHOUT LAW: HOW NEIGHBORS SETTLE DISPUTES 9 (1991).

64. *See* Lisa Bernstein, *Merchant Law in a Merchant Court: Rethinking the Code's Search for Immanent Business Norms*, 144 U. PA. L. REV. 1765, 1789 (1996); Lisa Bernstein, *Opting Out of the Legal System: Extralegal Contractual Relations in the Diamond Industry*, 21 J. LEGAL STUD. 115, 138 (1992); Lisa Bernstein, *Private Commercial Law in the Cotton Industry: Creating Cooperation Through Rules, Norms, and Institutions*, 99 MICH. L. REV. 1724, 1745 (2001).

65. *See* POSNER, *supra* note 1, at 18–19.

nations as vast and impersonal as the United States.[66] Their dependence on seemingly moribund sources of group commitment makes the prescriptions generated by these theories less behaviorally realistic and less politically feasible than those generated by the reciprocity theory, which—as the tax example illustrates—explains how and when individuals can be expected to cooperate even with strangers.

Just as important, however, their association with traditional sources of commitment makes these theories *morally* problematic. The cures that these theories propose for treating liberalism's major affliction risk vitiating liberalism's major accomplishment—the quieting of contentious social rivalries. The anxiety that the self-conscious regulation of social norms will infuse our political life with divisive "insider"-"outsider" distinctions furnishes the major source of opposition to the social norm project, both within and outside the academy. It deserves to be taken seriously by all social norm theorists. Yet Posner has little to say on this score.

Again, the reciprocity theory furnishes a better alternative. In contrast to the reputational signaling model as well as the other diverse social norm programs, the reciprocity theory doesn't rest on any particular set of social preconditions. To be sure, individuals will reciprocate with those whom they know and identify with; but they will also repay cooperation with cooperation in highly impersonal settings, such as labor markets and tax collection systems.[67] Interactions among individuals who deal repeatedly with one another will ground reciprocity dynamics; but so will the fleeting encounters among strangers, as laboratory experiments confirm.

The social malleability of reciprocity ultimately makes it more congenial to liberalism than are other social norm alternatives to the conventional theory of collective action. Because they don't depend on deep-seated group affinities, commitments founded on reciprocity can be built into the roles we inhabit as members of a liberal society. A white homophobic man in Mountain State cannot easily be made to display moral commitment to an Asian les-

66. *See generally* Francis Fukuyama, *Differing Disciplinary Perspectives on the Origins of Trust*, 81 B.U. L. REV. 479, 489–90 (2001).

67. *See* Kahan, *supra* note 35, at 11.

bian in Bay State when he appraises her as such. But if he can be made to think of her as a fellow *taxpayer*, *market participant*, or even *waste producer* he *can* be induced to reciprocate her cooperative behavior with cooperative behavior of his own.

Liberalism quiets social conflict by conditioning us to think of ourselves as unencumbered by divisive group commitments. The reciprocity model conserves this feature of liberalism, while correcting its tendency to make us think of ourselves as unencumbered by any moral commitments at all.

CONCLUSION

Eric Posner's *Law and Social Norms* is an impressive but ultimately flawed attempt to systematize the unruly elements of thought that pervade the growing literature on norms and law. His attempt to reduce norms to reputational signaling is behaviorally unrealistic. The prescriptions that his account generates are politically naïve. And the program that his model would commit us to as a means of correcting the deficiencies of liberal institutions is insufficiently respectful of the achievement of those institutions in quieting historic forms of social antagonism. Contending with Posner's theory will force social norm theorists to advance well beyond their current frameworks; however, in order for the social norms project to succeed, those theorists will have to advance far beyond Posner's as well.

***