

# The Logic of Reciprocity: Trust, Collective Action, and Law

Dan M. Kahan\*

## 1. Introduction

The *Logic of Collective Action*<sup>1</sup> has for decades supplied the logic of public policy analysis. In this pioneering application of public choice theory, Mancur Olson elegantly punctured the premise — shared by a diverse variety of political theories — that individuals can be expected to act consistently with the interest of the groups to which they belong. Absent externally imposed incentives, wealth-maximizing individuals, he argued, will rarely find it in their interest to contribute to goods that benefit the group as a whole, but rather will “free ride” on the contributions that other group members make. As a result, too few individuals will contribute sufficiently, and the well-being of the group will suffer.<sup>2</sup> These are the assumptions that dominate public policy analysis and ultimately public policy across a host of regulatory domains — from tax collection to environmental conservation, from street-level policing to policing of the internet.

But as a wealth of social science evidence (much of it appearing elsewhere in this volume) now makes clear, Olson’s *Logic* is false. In collective action settings, individuals adopt not a materially calculating posture but rather in a richer, more emotionally nuanced *reciprocal* one. When they perceive that others are behaving cooperatively, individuals are moved by honor, altruism, and like dispositions to contribute to public goods even without the inducement of material incentives. When, in contrast, they perceive that others are shirking or otherwise taking advantage of them, individuals are moved by resentment and pride to retaliate. In that circumstance, they will withhold beneficial forms of cooperation even if doing so exposes them to significant material disadvantage.<sup>3</sup>

---

\* Professor, Yale Law School.

<sup>1</sup> See Mancur Olson, *The Logic of Collective Action* (1965).

<sup>2</sup> See *id.* at 1-2.

<sup>3</sup> See Herbert Gintis, Samuel Bowles, Robert Boyd, and Ernst Fehr, *An Empirical and Analytical Examination of the Moral Sentiments*, chapter 1 of this volume; Ernst Fehr & Simon Gächter, *Reciprocity and Economics: The Economic Implications of Homo Reciprocans*, 42 *Euro. Econ. Rev.* 845 (1998); Ernst Fehr and Urs Fischbacher, *The Economics of Reciprocity*, in chapter 5, of this volume.

This behaviorally realistic picture of human motivation suggests not only an alternative account of when collective action problems will arise, but also an alternative program for solving (or simply avoiding) them through law. Whereas the conventional logic of collective action counsels the creation of appropriate external incentives, the new logic of reciprocity suggests the importance of promoting *trust*. Individuals who have faith in the willingness of others to contribute their fair share will voluntarily respond in kind. Spontaneous cooperation of this sort, moreover, breeds even more of the same, as individuals observe others contributing to public goods and are moved to reciprocate. In this self-sustaining atmosphere of trust, reliance on costly incentive schemes becomes less necessary. By the same token, individuals who lack faith in their fellows can be expected to resist contributing to public goods, inducing still others to withhold their cooperation as a means of retaliating. In this self-sustaining atmosphere of distrust, even strong (and costly) regulatory incentives are likely to be ineffective in promoting desirable behavior.

Indeed, such incentives may well undermine the conditions of trust necessary to hold collective action problems in check. Conspicuous rewards and punishments can imply that others *aren't* inclined to cooperate voluntarily, a message that predictably weakens individuals' commitment to contributing to public goods. In addition, incentive schemes tend to mask the extent to which individuals are inclined to contribute to public goods voluntarily, thereby weakening the tendency of observable cooperation to generate reciprocal cooperation by others. In short, manipulating material incentives may not only be an inefficient regulatory strategy for solving collective action problems; it may often be a self-defeating one.

This chapter will elaborate upon and apply these claims. It begins by distilling from the reciprocity literature a set of behavioral dynamics pertinent societal collective action problems. It then shows how these dynamics can be used to analyze and improve policymaking in various regulatory fields, with a particular emphasis on tax compliance, the siting of noxious facilities, and the policing of street crime.

## 2. The Logic of Reciprocity

Accepted for decades on a combination of faith and anecdote, the premises of the conventional theory of collective action have only recently been subjected to sustained and rigorous empirical examination. This research suggests an alternative "reciprocity theory" that differs

from the conventional position in four important respects, each of which merits specific attention.

**FIGURE 1**  
*Two Theories of Collective Action*

	Conventional Theory	Reciprocity Theory
Agents	Wealth maximizers	Emotional/moral reciprocators
Collective Behavior	Unique Equilibrium	Multiple Equilibria
Promoting Cooperation	Incentives	Trust
Variability of Preferences	Homogeneous	Heterogeneous

*Agents: Wealth Maximizers vs. Emotional/Moral Reciprocators*

The first pair of contrasting elements relates to the nature of individuals’ utility functions. The conventional theory assumes that individuals in collective action settings — ones that take the form of a standard prisoners’ dilemma — behave like wealth maximizers. That is, they refuse to contribute to collective goods and instead free-ride on the contributions made by others, who, as wealth maximizers, also contribute nothing. The reciprocity model, in contrast, sees individuals as moral and emotional reciprocators. Most persons think of themselves and want to be understood by others as cooperative and trustworthy and are thus perfectly willing to contribute their fair share to securing collective goods. But by the same token, most individuals are loath to being taken advantage of. Accordingly, if they perceive that most other individuals are shirking, they hold back too to avoid feeling exploited.

The reciprocity theory’s view of individual motivations is supported by a vast body of evidence. Much of it is experimental in nature. So-called “public goods” experiments — laboratory constructs designed to simulate collective action problems — have consistently shown that the willingness of individuals to make costly contributions to collective goods

is highly conditional on their perception that others are willing to do so.<sup>4</sup> Empirical studies of real-world behavior corroborate this finding. Individuals have been shown, for example, to reciprocate the disposition of others to give (or not) to charity,<sup>5</sup> to refrain (or not) from littering,<sup>6</sup> and to wait their turn (or not) in lines.<sup>7</sup> Indeed, individuals behave like reciprocators even in *markets*: econometric and other forms of field research, for example, suggest that when firms compensate their workers more generously workers reciprocate by voluntarily working harder.<sup>8</sup>

*Collective Behavior: Unique vs. Multiple Equilibria*

The second pair of contrasts concerns collective behavior. In typical collective action settings, the conventional theory treats defection or free-riding as the dominant strategy for every individual. Accordingly, that theory predicts a single collective behavioral equilibrium: universal noncooperation.

Under the reciprocity theory, in contrast, there is no “dominant” individual strategy. Individuals prefer to contribute if they believe others are inclined to contribute, but to free-ride if they believe that others are inclined to free-ride.

---

<sup>4</sup> See Fehr & Gächter, *supra* note 3.

<sup>5</sup> See Peter H. Reingen, *Test of a List Procedure for Inducing Compliance with a Request to Donate Money*, 67 J. Applied Psy. 110 (1982); see also Robert B. Cialdini, *Influence: Science and Practice* 96-97 (describing techniques used to create impressions of widespread charitable giving) (3d ed. 1993).

<sup>6</sup> See Elliot Aronson, *The Social Animal* 29-30 (7th ed. 1995); Robert B. Cialdini, Raymond R. Reno & Carl A. Kallgren, *A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places*, 58 J. Personality & Social Psy. 1015 (1990).

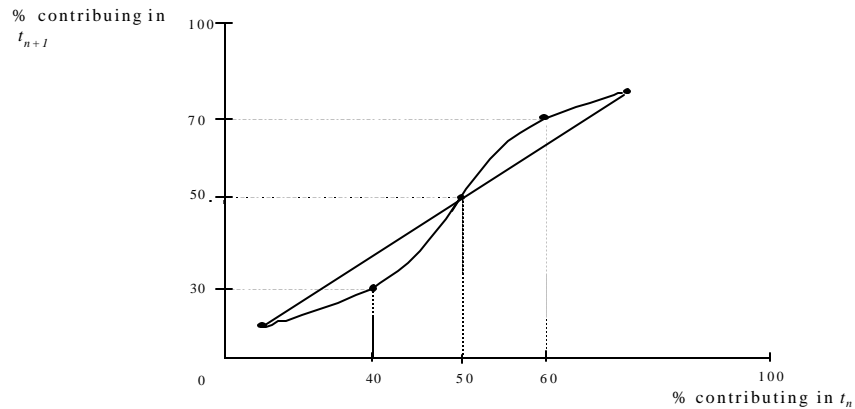
<sup>7</sup> See Stanley Milgram, Hilary James Liberty, Raymond Toldeo & Joyce Wackenhut, *Response to Intrusion into Waiting Lines*, 51 J. Personality & Social Psych. 683 (1986); Bernd Schmitt, Laurette Dubé, France Leclerc, *Intrusions into Waiting Lines: Does the Queue Constitute a Social System?*, 63 J. Personality & Social Psych. 806 (1992).

<sup>8</sup> See Gintis, et al., chapter 1 of this volume; George A. Akerlof, *Labor Contracts as Partial Gift Exchange*, 47 Q.J. Econ. 543 (1982); William Dickens & Lawrence Katz, *Inter-Industry Wage Differences and Theories of Wage Determination*, NBER Working Paper No. 2271, at 25-26 (1987); Lawrence Katz & Lawrence Summers, *Industry Rents: Evidence and Implications*, in *Brookings Papers on Economic Activity, Microeconomics* 209 (1989). See generally *Efficiency Wage Models of the Labor Market* (George A. Akerlof & Janet Yellen, eds., 1986).

Such interdependencies tend to generate patterns of collective behavior characterized by *multiple equilibria* punctuated by *tipping points*.<sup>9</sup> If, for whatever reason, some individuals conclude that those around them are inclined to contribute, they'll respond by contributing in kind, prompting still others to contribute, and so forth and so on until a

<sup>9</sup> These patterns can be illustrated graphically.

FIGURE 2  
Multiple Equilibria and Tipping Points



In this particular representation, there are three equilibria. One (selected arbitrarily for illustration) is around 50%: if participants in a collective action setting perceive that about half of the other participants are contributing in the period  $t_n$ , then about half will choose to contribute in the period  $t_{n+1}$ , which means that about that many will contribute in the period  $t_{n+2}$ , and so forth and so on. But this middle equilibrium is relatively unstable. If as a result of some exogenous shock, *more* than 50% are induced to contribute in  $t_n$  (say, 60%), then an even higher percentage than that will be willing to contribute in  $t_{n+1}$  (70%), leading a still higher percentage in  $t_{n+2}$ , and so forth and so on until contribution levels top out at the high-cooperation equilibrium at the upper right hand corner. Similarly, if for some reason *less* than 50% contribute in  $t_n$  (say, 40%), then an even smaller percentage will contribute in  $t_{n+1}$  (30%), leading to an even smaller contribution level in  $t_{n+2}$ , and so forth and so on until contributions bottom out at the low-cooperation equilibrium on the lower left hand corner. The corner equilibria, moreover, are relatively stable: exogenous shocks may result in temporary boosts or drops in contributions but unless they are big enough to push the contribution level back across the 50% tipping point, collective behavior will quickly settle back into the corner equilibrium from which it started. See generally Thomas C. Schelling, *Micromotives and Macrobehavior* (1978) (developing formal model of tipping points and feedback effects); Randal C. Picker, *Simple Games in a Complex World: A Generative Approach to the Adoption of Norms*, 64 U. Chi. L. Rev. 1225 (1997) (same).

highly cooperative state of affairs takes root. But if some individuals conclude that others are free-riding, then they will respond by free-riding, too, spurring others to do the same, and so forth and so on until a condition of mass noncooperation becomes the norm.

This dynamic, too, has been empirically documented. In multi-round public good experiments, for example, contribution levels tends to migrate steadily toward or away from the social optimum depending on whether subjects behaved relatively cooperative or noncooperatively early on.<sup>10</sup> Scholars have also documented that the incidence of littering, recycling, smoking in public, safe sex, and other types of behavior that affect collective welfare are likewise subject to feedback effects and multiple equilibria — generating dramatic variations in their incidence across space and over time.<sup>11</sup>

*Promoting Cooperation: Incentives vs. Trust*

The third contrast has to do with policy prescriptions. The conventional theory sees incentives as the solution to collective action problems: because wealth maximizers can't be counted on to contribute to public goods, they must be prodded to do so with either rewards or punishments that bring their individual interests into alignment with their collective ones.

The reciprocity theory suggests an alternative policy, *viz.*, the promotion of trust. If they can be made to believe that others are inclined to contribute to public goods, then individuals can be induced to contribute in turn, even without recourse to incentives. When permitted to communicate during play, for example, subjects in multi-round public goods experiments tend to assure one another that they'll contribute rather than free-ride. Although unenforceable, such assurances do in fact prompt larger contributions, which subjects quickly increase toward the social optimum as they observe others doing the same.<sup>12</sup> Face-to-face assurance-giving, in sum, conduces to trust, which in turn generates reciprocal cooperation.

---

<sup>10</sup> See generally Armin Falk and Urs Fischbacher, Modeling Strong Reciprocity, Chapter 6 of this volume.

<sup>11</sup> See, e.g., Robert Cooter, *Normative Failure Theory of Law*, 82 Cornell L. Rev. 947, 976-77 (1997); Timur Kuran & Cass R. Sunstein, *Availability Cascades and Risk Regulation*, 51 Stan. L. Rev. 683, 688-89, 746 (1998); Cass R. Sunstein, *On the Expressive Function of Law*, 144 U. Pa. L. Rev. 2021, 2032-36 (1996).

<sup>12</sup> See Ledyard, *supra* note 10, at 156-68; Elinor Ostrom, *Collective Action and the Evolution of Social Norms*, 14 J. Econ. Perspectives 137, 146 (2000)

Indeed, field and laboratory research suggests that incentives, far from solving collective action problems, can sometimes actually magnify them by dissipating trust. The simple existence of an incentive scheme can be seen as a *cue* that other individuals are not inclined to cooperate voluntarily: if they were, incentives would be unnecessary. This inference can in turn trigger a reciprocal disposition to withhold voluntary cooperation thereby undercutting, if not wholly displacing, the force of the incentive. In addition, the existence of incentives can *mask* voluntary contributions to public goods, thereby diluting the power of such contributions to trigger reciprocal cooperation. Relatedly, incentives can *crowd out* dispositions such as altruism by extinguishing the opportunity of individuals to demonstrate (to themselves and to others) that they are willing to sacrifice material gain for the public good. And if for any of these reasons, the advent of a material incentive for this reason induces even a few individuals to contribute less to a public good, moreover, reciprocity dynamics will induce still others to contribute less, thereby inducing others to do the same, and so forth and so on until collective behavior settles into a new, noncooperative equilibrium — one that is likely to be impervious to the subsequent removal of material incentives.<sup>13</sup>

It would be a mistake, though, to conclude that material incentives invariably diminish trust. They are most likely to have that effect, research suggests, when individuals start out with the belief that most other individuals are inclined to contribute to some public good voluntarily; that's when the advent of material incentives creates the greatest risk of adverse cueing, masking, and crowding out. But things are likely to be different if individuals start out with the belief that most other individuals are inclined to shirk or free-ride. In that case, the advent of a credible reward or penalty can work — not just by changing individuals material incentives but by changing in a positive way their impression of the willingness of other individuals to behave cooperatively rather than noncooperatively in a collective action setting.

---

<sup>13</sup> See Elinor Ostrom, *Policies that Crowd Out Reciprocity and Collective Action*, Chapter 9 of this volume. See generally Bruno S. Frey, *Not Just for the Money: An Economic Theory of Personal Motivation* (1997); Uri Gneezy & Aldo Rustichini, *A Fine Is a Price*, 29 *J. Legal Stud.* 1 (2000) (finding that fine increased rather than decreased abuse of day care center rules by parents); Uri Gneezy & Aldo Rustichini, *Pay Enough or Don't Pay at All* (unpublished manuscript, April 1999) (finding that incentives decreased rather than increased performance of individuals soliciting charitable donations); Richard M. Titmuss, *The Gift Relationship: From Human Blood to Social Policy* (1971) (finding incentives suppress donation of blood); Bruno S. Frey & Reto Jegen, *Motivation Crowding Theory: A Survey of Empirical Evidence*, *J. Econ. Surveys* (forthcoming).

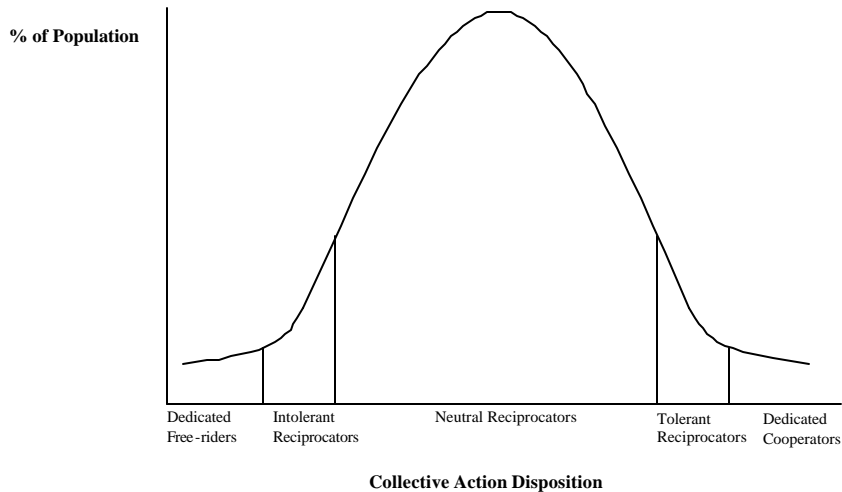
An example is the power of higher-than-average wages to elicit higher-than-average productivity in the workplace. Workers naturally suspect their firms of being unwilling to share a fair portion of the surplus generated by the workers' labor. But when a firm offers workers a wage that exceeds the industry average, workers are likely to infer that that particular firm *is* willing to divide the surplus fairly; they therefore respond by voluntarily working more productively, which inclines firms to maintain or even raise their wages. The result is a self-sustaining form of reciprocal cooperation that obviates the need for costly performance monitoring regimes.

*Variability of Preferences: Homogeneous vs. Heterogeneous*

Finally, the conventional theory and the reciprocity theory differ on the variability of preferences across individuals. The conventional theory imagines that the disposition to free-ride in collective action settings is relatively uniform. In contrast, the evidence on which the reciprocity theory rests suggests that the disposition to cooperate varies. In public goods experiments that generate multiple equilibria, for example, neither universal cooperation nor universal defection is the final resting point.

It makes more sense, then, to envision a distribution of cooperative dispositions across the population.

**FIGURE 3**  
*Heterogeneity of Collective Action Dispositions*



Some, relatively small fraction of the population (consisting, perhaps, of



those who've been trained in neo-classical economics) consists of committed free-riders, who shirk no matter what anyone else does, and another small fraction (maybe those who've read too much Kantian moral philosophy) of dedicated cooperators, who contribute no matter what. But most individuals are reciprocators, who cooperate conditionally on the willingness of others to contribute. Moreover, some reciprocators are relatively intolerant: they bolt as soon as they observe anyone else free-riding. Others are relatively tolerant, continuing to contribute even in the face of what they see as a relatively modest degree of defection. And a great many more — call them the neutral reciprocators — fall somewhere in between.

Under these circumstances, individuals are unlikely fully to overcome collective action problems through reciprocity dynamics alone. No matter how cooperative the behavior of others, the committed free-riders will always free-ride if they can get away with it. Indeed, their shirking could easily provoke non-cooperative behavior by the less tolerant reciprocators, whose defection in turn risks inducing the neutral reciprocators to abandon ship, thereby prompting even the tolerant reciprocators to throw in the towel, and so forth and so on. If this unfortunate chain reaction takes place, a state of affairs once characterized by a reasonably high degree of cooperation could tip decisively toward a noncooperative equilibrium in which only the angelic, Kantian, unconditional cooperators are left contributing (probably futilely) to the relevant public good.

Maximum cooperation, then, probably requires that reciprocity dynamics be supplemented with *appropriately tailored* incentives, most likely in the form of penalties aimed specifically at persistent free-riders. Although trust and reciprocity elicit cooperation from most players, some coercive mechanism remains necessary for the small population of dedicated free-riders, who continue to hold out in the face of widespread spontaneous cooperation, thereby depressing the contributions made by some, relatively unforgiving reciprocators. In the face of a credible penalty, however, the committed free-riders fall into line. The existence of such penalties in turn assures the less tolerant reciprocators that their cooperation won't make them into chumps; they thus continue to cooperate, less out of material interest than out of positive reciprocal motivations. And because the less tolerant reciprocators contribute, so do the neutral and tolerant reciprocators, generating an equilibrium of near-universal cooperation. Again, these dynamics are borne out by empirical evidence,

particularly ones that allow subjects in public goods experiments to retaliate against defectors.<sup>14</sup>

The uneven effect of penalties in promoting and dissipating trust calls attention to the *expressive* dimension of incentives. Incentives do more than affect individuals' calculations of the costs and benefits of particular forms of conduct; they also shape their impressions of the attitudes and intentions of those around them.<sup>15</sup> Laboratory and real-world schemes that use *generally applicable incentives* convey the message that noncooperation is the norm, and thus stifle the reciprocal motivations of even neutral reciprocators, whose defection predictably spills over onto even the most forgiving ones. Targeted retaliation, in contrast, conveys a very different message. Because all individuals are aware from social experience that there are some committed free-riders out there, no one is surprised or disappointed to see penalties aimed at *those types*; accordingly, such penalties don't create the cueing, masking, or crowding out effects associated with more generalized incentive regimes. On the contrary, penalties understood to be necessitated only by the existence of committed free-riders have a trust-enhancing effect, for they imply that *most* individuals are *not* inclined to shirk. Targeted retaliation works, in sum, because it simultaneously coerces dedicated free-riders, calms unforgiving reciprocators, and *avoids* confusing or demoralizing neutral and forgiving reciprocators.

### 3. Tax Compliance

Tax compliance is the consummate collective action from a public policy point of view. Society collects taxes to finance a variety of goods — from education to highways to national defense — that benefit its members collectively. Nevertheless, it is in the individual material interest of every citizen to free-ride on her fellow citizens' contributions to these goods while withholding any contribution of her own. Accordingly, the conventional theory predicts that individuals, as wealth maximizers, will evade their taxes, unless furnished with incentives — in the form of

---

<sup>14</sup> See Gintis, et al., chapter 1 of this volume; Fehr & Fischbacher, chapter 5 of this volume; Ernst Fehr & Simon Gächter, *Cooperation and Punishment in Public Goods Experiments*, 90 Am. Econ. Rev. 980 (2000).

<sup>15</sup> See generally Dan M. Kahan, *Social Influence, Social Meaning, and Deterrence*, 83 Va. L. Rev. 349 (1997).

threatened penalties — that make the expected return from evasion smaller than the expected return from compliance.<sup>16</sup>

This account of tax evasion is embarrassingly ill-supported by empirical evidence. Econometric studies have concluded that the expected penalty for evasion explains little if any of the variation in compliance across space or over time.<sup>17</sup> Survey measures also find only very modest correlation between reported compliance and individuals' subjective perception of the expected penalty for evasion.<sup>18</sup> Finally, laboratory experiments that simulate the decision to evade suggest that probability and severity of detection can influence individual decisions to evade, but only when they are set at levels far in excess of those associated with actual policies.<sup>19</sup>

What explains a lot more, empirical research suggests, is a complex of factual beliefs and emotional dispositions. Thus, an individual's perception of the extent of evasion is a powerful predictor of compliance behavior: the higher an individual believes the rate of tax-cheating to be, the more likely he or she is to cheat too.<sup>20</sup> So too are the prospect of shame (or potential stigma) and guilt. The more likely an individual believes it is that she will be condemned by others should she be caught, the more likely she is to refrain from evading. By the same token, the more regret

---

<sup>16</sup> See generally Micale G Allingham & Agnar Sandomo, *Income Tax Evasion: A Theoretical Analysis*, 1 J. Pub. Econ. 323 (1972).

<sup>17</sup> See *id.* at 842; Frank A. Cowell, *Cheating the Government: The Economics of Evasion* 74 (1990); Steven Klepper & Daniel Nagin, *The Criminal Deterrence Literature: Implications for Research on Taxpayer Compliance*, in 2 *Taxpayer Compliance* 126, 142 (J. Roth and J. T. Scholz ed., 1989).

<sup>18</sup> See, e.g., Harold G. Grasmick & Wilbur J. Scott, *Tax Evasion and Mechanisms of Social Control: A Comparison with Grand and Petty Theft*, 2 J. Econ. Psych. 213, 225 & 226 tbl. 2 (1982).

<sup>19</sup> See James Andreoni, Brian Erard & Jonathan Feinstein, *Tax Compliance*, 36 J. Econ. Lit. 818, 841 (1998).

<sup>20</sup> See Robert B. Cialdini, *Social Motivations to Comply: Norms, Values, and Principles*, in 2 *Taxpayer Compliance* 215 (J. A. Roth and J. T. Scholz ed., 1989); James P.F. Gordon, *Individual Morality and Reputations Costs as Deterrents to Tax Evasion*, 33 Euro. Econ. Rev. 797 (1989); Klepper & Nagin, *supra* note 17, at 144; Steven M. Sheffrin & Robert K. Triest, *Can Brute Deterrence Backfire? Perceptions and Attitudes in Taxpayer Compliance*, in *Why People Pay Taxes* 193 (J. Slemrod ed., 1992).

or remorse that an individual believes she'd experience for engaging in evasion, the less likely she is to engage in that crime.<sup>21</sup>

These are exactly the factors one would expect to influence tax compliance were individuals behaving like moral and emotional reciprocators. A strong reciprocator wants to understand herself and be understood by others as fair, but she loathes 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 comes into play. *If* most other individuals seem to be paying their taxes, then evasion will provoke either guilt, shame, or both in the reciprocator who covets the respect of others and of himself. If, in contrast, most individuals appear to be evading, then complying *won't* make her feel guilty or ashamed at all; it will make her feel like a sucker.

This interpretation of the data is confirmed by an experiment that tested how the 1986 Tax Reform Act affected compliance levels.<sup>22</sup> One hypothesis, suggested by the conventional theory, was that individuals would become more or less willing to evade depending on whether the Act had increased or decreased their relative tax burden. The study found no such correlation. What *did* shift patterns of compliance, the researchers found, were the types of interactions that individuals had with other taxpayers in the months leading up to the reform: those who encountered others who expressed a positive attitude toward, and commitment to complying with, the Tax Reform Act displayed greater commitment to complying with it themselves, whereas those who encountered others who expressed *negative* attitudes displayed less commitment.<sup>23</sup> This effect, moreover, was explained completely by variation in the shame and guilt that the two groups of taxpayers anticipated for failing to pay their taxes.<sup>24</sup> In other words, as moral and emotional reciprocators, these individuals naturally felt guilt and shame for failing to contribute to the public good of tax compliance in proportion to their perception that others were or were not contributing.

---

<sup>21</sup> See, e.g., Grasmick & Scott, *supra* note 18, at 226 & tbl. 4; Wilbur J. Scott & Harold G. Grasmick, Deterrence and Income Tax Cheating: Testing Interaction Hypotheses in Utilitarian Theories, 17 J. Applied Behavioral Sci. 395, 403 tbl. 1 (1981).

<sup>22</sup> 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* 9 (Joel Slemrod ed. 1992).

<sup>23</sup> See *id.* at 29-30.

<sup>24</sup> See *id.*

The conventional theory of collective action is just as weak at explaining variance in tax compliance across nations as it is in explaining compliance across individuals. Tax compliance rates vary dramatically across nations. Essentially none of this variance, however, can be explained by differences in the expected penalty for evasion. More important, researchers have concluded, are differences in public attitudes toward tax laws. In some nations (including the United States), individuals tend to view paying their taxes as an important civic obligation, and are highly motivated to pay for that reason. In other nations (including many in Western Europe), individuals regard tax obligations much more casually (akin, say, to traffic regulations in the United States), and display no particular moral aversion to evading them if they feel they can safely do so.<sup>25</sup>

Varying national “tax cultures” of this sort are perfectly understandable under 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 material payoffs associated with cooperating or defecting. 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, 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.<sup>26</sup>

The empirical evidence also bears out the trust theory’s anxiety about the effect of self-defeating material incentives. Experimental evidence suggests that when taxpayers are exposed to information highlighting the penalties for evasion, they respond in much the same way that subjects in public goods experiments do when furnished with generalized

---

<sup>25</sup> See James Alm, Isabel Sanchez & Ana De Juan, *Economic and Noneconomic Factors in Tax Compliance*, 48 KYKLOS 3 (1995); Cowell, *supra* note 17, at 102-03.

<sup>26</sup> See Steven M. Sheffrin & Robert K. Triest, *Can Brute Deterrence Backfire? Perceptions and Attitudes in Taxpayer Compliance*, in *Why People Pay Taxes* 193, 194-95 (J. Slemrod ed., 1992) (suggesting interdependence of taxpayer decisionmaking should generate multiple behavioral equilibria); see also Cowell, *supra* note 17, 112-13 (developing theoretical model predicting multiple compliance equilibria based on interdependence of taxpayers’ decisions to evade).

material incentives to contribute — namely, by contributing *less*.<sup>27</sup> Researchers have also found that the highly politicized auditing campaigns tend to provoke a *higher* incidence of tax cheating rather than a lower one.<sup>28</sup>

The mechanism for these effects appears to be social cueing. When government engages in dramatic gestures to make individuals aware that the penalties for tax evasion are being increased, it also causes individuals to infer that more taxpayers than they thought are choosing to cheat. This *distrust* of one's neighbors triggers a reciprocal motive to evade, which dominates the greater material incentive to comply associated with the higher than expected penalty.<sup>29</sup>

Is there a way for tax enforcers to *bolster* taxpayers' trust in one another? One policy that seems to do that is simply to advise citizens that the vast majority of taxpayers *are* in fact complying. In a study sponsored by the Minnesota Department of Revenue, 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. Those individuals thereafter reported more income and claimed fewer deductions than did individuals in a control group. This is exactly what the phenomenon of reciprocity would predict: when they learn that others are in fact disposed to contribute their fair share, individual taxpayers, just like individuals in public good experiments, cooperatively respond in kind. Likewise consistent with the reciprocity theory — and at odds with the conventional economic one — the Minnesota study found that individuals advised of high compliance rates paid more tax than did individuals who received letters advising them that their returns would be subject to a greater rate of auditing!<sup>30</sup>

Another policy that appears to promote trust and hence bolster reciprocal cooperation is the enactment of popular reforms. As the study of the 1986 Tax Reform Act demonstrates, such reforms promote the expression of positive views toward the law. When they are exposed to

---

<sup>27</sup> See Richard D. Schwartz & Soya Orleans, *On Legal Sanctions*, 34 U. Chi. L. Rev. 274, 298 (1967).

<sup>28</sup> See Steven M. Sheffrin & Robert K. Triest, *Can Brute Deterrence Backfire? Perceptions and Attitudes in Taxpayer Compliance*, in *Why People Pay Taxes* 193, 211-13 (J. Slemrod ed., 1992).

<sup>29</sup> See *id.*

<sup>30</sup> See Stephen Coleman, *The Minnesota Income Tax Compliance Experiment: State Tax Results* (1996).

these views, individuals infer that others are inclined to comply. That conclusion in turn triggers the disposition to reciprocate in kind. In effect, the enactment of popular reforms generates an environment of face-to-face assurance giving that builds trust, and a resulting disposition to cooperate, in much the same way that discussion does in public goods experiments.

The contribution that reciprocity makes to tax compliance doesn't imply that the IRS should disavow punishments for evasion altogether. That would be foolhardy because of the variability of individual dispositions to cooperate in collective action settings. With no risk of punishment, evasion would become commonplace among dedicated cheaters, whose defections would in turn unleash a contagious form of demoralization among the vast run of reciprocity-minded taxpayers.

The difference between effective incentives and ineffective ones, experimental and other empirical data suggest, lies in the social meanings that they express. Enforcers should therefore carefully select cases to nourish the perception that evaders are deviants, not normal citizens.<sup>31</sup> It is already common belief that a certain number of individuals of exceptional venality will evade even when nearly all the rest of us are complying. The existence of coercive incentives understood to be aimed at *those* persons, then, doesn't dispel trust; on the contrary, it helps to assure the honest multitudes that they are not being exploited when they choose to pay their taxes. A model case, in this sense, was the tax-fraud prosecution of hotel-magnate Leona Helmsley, who expressed open contempt for income taxes as something that "only the little people pay."<sup>32</sup>

In addition, officials should always juxtapose trust-enhancing information with penalties. Auditing crack downs and other high-profile modes of enforcement risk backfiring, the evidence suggests, because they function as a *cue* that evasion is widespread. To counteract this inference, enforcers should be sure that the *good news* that the vast majority of citizens voluntarily comply always gets at least equal billing with the bad news that a small minority don't. They should take advantage of the attention that high-profile prosecutions naturally attract to publicize positive information akin to that shown to generate even higher rates of compliance in the Minnesota Tax Experiment.

---

<sup>31</sup> See Cialdini, *supra* note 20, at 215.

<sup>32</sup> See *The Wicked Witch Who Has Poisoned the Big Apple*, Times (London), Sept. 3, 1989 (" 'She deserves everything she gets, she's scum,' said one of hundreds of people who waited outside the federal courthouse in Manhattan on Wednesday to jeer at Leona.").

Unfortunately, they often do just the opposite. Competing with other agencies and programs for appropriations, the IRS routinely exaggerates the inadequacy of its own enforcement powers and the resulting extent of evasion.<sup>33</sup> Usually timed to be reported on the media the week before personal income taxes are due, IRS-generated stories of the agency's own inefficacy in enforcing the law predictably generates resentment in those who routinely obey.<sup>34</sup> "Are You a Chump?" a *Forbes* magazine cover story asked its tax-paying readers as the magazine reported on the supposed decimation of the IRS' enforcement capacity.<sup>35</sup>

The United States, in truth, enjoys a relatively high compliance rate. But that hardly means that things can't be made worse. Like other high-cooperation equilibria sustained by reciprocity dynamics, the disposition of Americans voluntarily to pay their taxes can be "tipped." If by rattling its saber one day and pleading poverty the next, the IRS succeeds in inducing enough taxpayers to believe that cheating is indeed widespread, the result could be a self-reinforcing wave of evasion. The result could be a new, low-cooperation equilibrium that, as the durability of Europe's disobedient tax culture attests, can be very difficult to reverse. Ironically, by embracing the conventional-theory strategy of "incentives, incentives, and more incentives," the IRS risks making tax compliance into exactly the type of intractable collective action problem that the conventional theory envisions it to be.

#### 4. "Not in My Backyard"

Various types of public facilities — including highways, airports, prisons, hazardous waste dumps, and the like — impose disproportionate burdens (noise, perceived physical danger, health risks) on persons who reside near them. Accordingly, even when they recognize the benefits of these facilities for society at large, individuals often resist the siting of

---

<sup>33</sup> See, e.g., David Cay Johnston, *A Smaller I.R.S. Gives Up On Billions in Back Taxes*, N.Y. Times, Apr. 13, 2001, at A1.

<sup>34</sup> See, e.g., Tom Brazaitis, *Wimpy IRS Emboldens Cheats*, Plain Dealer (Cleveland, OH), Apr. 18, 2001 at 11B; Amy Feldman & Joan Caplin, *Should You Cheat on Your Taxes?*, Money, Apr. 2001, at 108.

<sup>35</sup> Janet Novack, *Are You a Chump?*, Forbes, Mar. 5, 2001, at 122.



them within their own communities, a phenomenon that political scientists refer to as the “not in my backyard” phenomenon or “NIMBY.”<sup>36</sup>

The conventional theory of collective action sees NIMBY as another expression of individuals’ propensity to withhold costly contributions to public goods and instead to free ride on the contributions of others. Accordingly, the standard model proposes an incentives-based solution: that the communities best-situated to host a particular facility be compensated for the burden associated with it, presumably out of the proceeds of a tax imposed on the individuals who benefit from the facility but who reside elsewhere.<sup>37</sup>

This strategy, however, has an unimpressive track record. In the 20 years since Massachusetts enacted a widely lauded compensation scheme, not a single community has accepted — or been forced to accept — a facility siting.<sup>38</sup> The results have been the same in numerous other states and Canadian provinces that have tried to induce siting with compensation.<sup>39</sup>

Indeed, there is evidence that compensation schemes at least sometimes make the NIMBY problem *worse*. According to some studies, residents often bridle at “compensation offers . . . as attempts to buy them off or bribe them.”<sup>40</sup> The potential of incentives to backfire in this way has been confirmed experimentally by Swiss economists Bruno Frey and Felix Oberholzer-Gee, who showed that a compensation offer dramatically reduced (from just over 50% to less than 25%) the number of laboratory subjects willing to assent to the siting of a nuclear waste storage facility in their community.<sup>41</sup>

---

<sup>36</sup> See generally Don Munton, *Introduction: The NIMBY Problem and Approaches to Facility Siting*, in *Hazardous Waste Siting and Democratic Choice 1* (D. Munton ed., 1996); Barry G. Rabe, *Beyond NIMBY: Hazardous Waste Siting in Canada and the United States 1-2* (1994).

<sup>37</sup> The classic statement of this analysis is Michael O’Hare, “*Not on My Block You Don’t*”: *Facility Siting and the Strategic Importance of Compensation*, 25 *Pub. Pol.* 407 (1977).

<sup>38</sup> See Kent E. Portney, *Siting Hazardous Waste Treatment Facilities* 28 (1991); Rabe, *supra* note 36, at 36-37.

<sup>39</sup> See *id.* at 39-44.

<sup>40</sup> Munton, *supra* note 36, at 17.

<sup>41</sup> See Frey, *supra* note 13, 69-75.

It would be a mistake, however, to conclude that compensation schemes *never* work. At least some opinion studies have shown that offers of compensation can significantly increase willingness to accept the siting of a noxious facility.<sup>42</sup> Moreover, compensation in one form or another *has* nearly always been a part of the successful waste-facility siting efforts in the United States and Canada in recent decades.<sup>43</sup>

While failures predominate, it's fair to conclude that "studies show a high degree of variability in the ability of compensation to change public opinion" toward siting.<sup>44</sup> But precisely because they are *not* uniformly positive, these results furnish little support for the conventional theory's account of NIMBY. Clearly, something more than the weighing of material costs and benefits is going on when communities decide whether to resist or to accept noxious facilities.

That something more, opinion analyses suggest, is the moral and emotional reaction of residents to siting proposals. Individuals who interpret the decision to impose a site on their community as signifying the low social status of its residents — who believe that they are being "dumped on," symbolically as well as literally — are more likely to resist.<sup>45</sup> Those who distrust government institutions generally also are less likely to tolerate the siting of a noxious facility in their vicinity,<sup>46</sup> as are those who believe that societal benefits and burdens in general, and the burdens associated with the facility in question in particular, are being distributed ineq-

---

<sup>42</sup> See Howard Kunreuther & Doug Easterling, *The Role of Compensation in Siting Hazardous Facilities*, 15 *J. Policy Analysis & Management* 601, 605-06 (1996); Howard Kunreuther, Douglas Easterling, William Desvousges & Paul Slovic, *Public Attitudes Toward Siting a High-Level Nuclear Waste Repository in Nevada*, 10 *Risk Analysis* 469, 480 (1990).

<sup>43</sup> See Munton, *supra* note 36, at 16; Douglas J. Lober, *Beyond NIMBY: Public Attitudes and Behavior and Waste Facility Siting Policy* 124-25 (Ph.D. Dissertation, Yale University, School of Forestry & Environ. Stud., 1993).

<sup>44</sup> See at Kunreuther & Easterling, *supra* note 42, at 605.

<sup>45</sup> Lober, *supra* note 43, at 120; see also Kunreuther *et al.*, *supra* note 42, at 470; see also Paul Slovic, M. Layman, N. Kraus, James Flynn, J. Chalmers & G. Gesell, *Perceived Risk, Stigma, and Potential Economic Impacts of High-Level Nuclear Waste Repository in Nevada*, in *Risk, Media, and Stigma* (James Flynn, Paul Slovic & Howard Kunreuther eds., 2001).

<sup>46</sup> See Robin Gregory, Howard Kunreuther, Doug Easterling & Ken Richards, *Incentive Policies to Site Hazardous Waste Facilities*, 11 *Risk Analysis* 667, 672 (1991); Kunreuther *et al.*, *supra* note 42, at 472; Lober, *supra* note 43, at 140-42.

uitably.<sup>47</sup> The perception that the racial composition of the community is playing a role in that process can create intense opposition in minority communities, which historically have been least able to muster the political resources necessary to resist forced sitings.<sup>48</sup>

These are the sorts of factors that one would expect to influence the reactions of individuals who behave like moral and emotional reciprocators with respect to civic obligations. When called upon to accept risks or inconveniences in the interest of the public good, individuals who believe that societal benefits and burdens are being inequitably distributed by fundamentally unjust political institutions unsurprisingly answer, “No.”

Reciprocal motivations also explain another factor relevant to acceptance of toxic waste facilities: the *origin* of the wastes. A wealth-maximization model suggests that waste source should be irrelevant: home-grown wastes are every bit as hazardous as out-of-town ones. But in fact, individuals are much more likely to accept disposal facilities for wastes produced locally.<sup>49</sup> This makes sense insofar individuals are likely to accept a waste disposal facility in a spirit of positive reciprocation when they understand the waste to have been generated by beneficial local activities.

The uneven effect of compensation schemes also conforms to the logic of reciprocity, which implies that the effect of incentives in dissipating or promoting trust depends critically on citizens’ moral and emotional priors. Imagine a society whose citizens begin with the belief that societal burdens *are* being equitably distributed through a *just* political process. We might expect those individuals, as reciprocators, to be relatively accepting of the siting of noxious facilities in their community. But if authorities try to purchase acceptance with incentives, these same individuals might revise their views, inferring that other communities must in fact be *unwilling* to accept such impositions voluntarily. As a result of this perverse cueing effect, the NIMBY phenomenon will *grow* in strength, as individuals reciprocate the perceived resistance to such facilities by strengthening their own resistance to them.

This reaction plausibly explains the results in the Frey/Oberholzer-Gee experiment. Homogeneous, democratic, and small, Switzerland has an admirable history of resolving disputes over the allocation of societal benefits and burdens through a fair process of deliberative give-and-take.

---

<sup>47</sup> See Kunreuther & Easterling, *supra* note 42, at 601-02; Lober, *supra* note 43, at 145.

<sup>48</sup> See *id.* at 145; Rabe, *supra* note 36, at 21.

<sup>49</sup> See Lober, *supra* note 43, at 126; Rabe, *supra* note 36, at 44.

The Swiss subjects in the experiment therefore interpreted the offer of a cash payment as evidence that the norm of mutual accommodation had broken down in the case of nuclear wastes and became predictably indignant at attempts to buy their assent to a risk that others refused to endure.

But now imagine the perhaps more typical U.S. or Canadian case of a community whose residents start off with the belief that society's resources are being *inequitably* distributed as a result of a fundamentally *unjust* political system. As reciprocators, they are likely to resist the nearby siting of a noxious facility. Yet in that climate, there is at least some potential for compensation to work: not only does compensation help to offset the material inconveniences or risks associated with the facility; the very offering of it conveys a degree of respect that previously had been lacking in their political life.

Case studies suggest that this result is most likely when incentives are part of a negotiated, "bottom-up" siting regime rather than a centrally administered "top-down" one.<sup>50</sup> Even with compensation, the imposition of a site by a centralized bureaucracy is likely to provoke negative reciprocal motivations. The authority of administrators to dictate the site location suggests that others are unwilling to accept the facility voluntarily, a signal that is reinforced by the need to offer compensation. When voluntary acceptance is solicited, however, communities that historically have been disadvantaged are likely to feel respected and empowered; compensation is no longer seen as degrading. In addition, the process of negotiation is likely to create a climate akin to the face-to-face discussions in public goods games: when they are able to discuss the situation with remote political authorities, and are granted veto power, local communities are likely to be assured that others are willing to contribute fair share to dealing with the problem. Accordingly, they reciprocate positively by showing greater receptivity to placement of the facility.

These effects, case studies suggest, feed on each other, generating multiple behavioral equilibria. Again, in Massachusetts, which enacted a top-down, dictate-plus-compensation regime in the 1980s, one community after another fought off attempts to site hazardous waste facilities within their borders, whereas in Wisconsin, which has a bottom-up, negotiated-compensation scheme, a succession of communities have come forward

---

<sup>50</sup> See generally See Kunreuther & Easterling, *supra* note 42, at 618; Munton, *supra* note 36, at 19-20; Rabe, *supra* note 36, at 59.

to accept such facilities.<sup>51</sup> Provinces in Western Canada have had similar strings of successes with the negotiated-compensation strategy.<sup>52</sup>

The key to solving NIMBY, in short, is trust. Various sources of evidence suggest that individuals *can* be made receptive to the siting of noxious facilities in their communities *if* they can be made to believe that society is committed to treating their interests with respect. Appropriately structured bottom-up, negotiated-compensation schemes — ones framed to emphasize respect for the interests and autonomy of prospective host communities — are one way to reverse deep-seated resentments and thus excite a reciprocal openness to siting decisions. If individuals *can't* be made to believe that the burden of accepting a noxious facility is being fairly reciprocated either in kind or by like sacrifices, the current of resentment that fuels NIMBY will be difficult to reverse, even with financial incentives.

## 5. Street Crime

The conventional theory sees crime prevention as just another collective action problem. As a society, we are all better off when we universally refrain from theft and like forms of predation. But as individuals, each one of us is better off free-riding on whatever restraint our neighbors display while engaging in as much looting and pillaging as possible. Public order is, in short, a public good, one that will always be in short supply if individuals are left to their own devices.

If this is how one thinks of the problem of crime, then the obvious solution is to create incentives that bring individual interests into alignment with collective ones. Hence, the threat of punishments for those who break the law.

The conventional theory of collective action thus naturally gives rise to the law-enforcement strategy of deterrence, which can be neatly formalized in terms first proposed by Bentham<sup>53</sup> and later refined by

---

<sup>51</sup> See Kunreuther & Easterling, *supra* note 42, at 618; Lober, *supra* note 43, at 222-23.

<sup>52</sup> See Geoffrey Castle & Don Munton, *Voluntary Siting of Hazardous Waste Facilities in Western Canada*, in *Hazardous Waste Siting and Democratic Choice* 56-57 (D. Munton ed., 1996); Rabe, *supra* note 36, at 61-81.

<sup>53</sup> See Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation*, reprinted in *The Utilitarians* (1961).

Becker.<sup>54</sup> As wealth maximizers, individuals, on this theory, commit crime when the *gain*,  $G$ , is greater than the *expected punishment*, which is equal to product of the specified penalty,  $P$ , and the certainty,  $C$ , that it will be imposed. Thus, crime is deterred when  $P \times C > G$ .<sup>55</sup>

Of course, it is efficient or collectively wealth-maximizing to deter crime only if the social cost of  $P \times C$  is less than the social losses associated with the crimes that  $P \times C$  deters. Accordingly, society must be attentive to the cost of various  $P \times C$  pairings. This attentiveness generally favors severity over certainty, since maintaining a high likelihood of detection and conviction ( $C$ ) requires a continuing investment in police officers, judges, prosecutors, public defenders, etc., whereas a high level of punishment ( $P$ ) — assuming it deters and thus doesn't have to be imposed all that often — won't cost much to implement and will allow society to economize on the various components of law enforcement.<sup>56</sup>

This turns out to be a fair summary of the guiding philosophy of American criminal law enforcement in the last twenty-five years — the results of which do little to vindicate the wisdom of the conventional theory. Variance in the severity of punishment has consistently been shown to explain little, if any, of the variance in incidences of robbery, burglary, homicide, drug dealing and other street crimes across place and time. Certainty of conviction makes a difference, although a relatively small one.<sup>57</sup>

What matters much more are a diverse collection of social conditions and public attitudes. Thus, communities characterized by low “social organization” — as measured by the quality and vitality of voluntary civic associations — tend to have more crime.<sup>58</sup> So do ones in which institutions lack “legitimacy,” as measured by the willingness of individuals to view the decisions of law-makers and -enforcers as intrinsically entitled

---

<sup>54</sup> See Gary Becker, *Crime and Punishment: An Economic Approach*, 76 J. of Pol. Econ. 169 (1968).

<sup>55</sup> See *id.*

<sup>56</sup> See *id.*; Richard Posner, *An Economic Theory of Crime*, 85 Colum. L. Rev. 1193 (1985).

<sup>57</sup> See generally Daniel Nagin, *Criminal Deterrence Research at the Outset of the Twenty-First Century*, 23 Crim. & J. 1 (1998).

<sup>58</sup> See Robert J. Sampson, Stephen W. Raudenbush & Felton Earls, *Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy*, 277 Science 918 (1997).

to deference.<sup>59</sup> “Social influence” — the tendency of individuals to conform their behavior to those around them — also contributes to the incidence of crime, generating multiple crime-rate equilibria independent of the expected penalty for law-breaking.<sup>60</sup>

Where these factors are conducive to criminality, many individuals will break the law *notwithstanding* very severe penalties. Indeed, there is reason to believe that severe penalties can deleteriously affect the attitudes and social conditions that lead to crime: massive incarceration, particularly when concentrated on minority, inner-city communities, disrupts social organization and taxes institutional legitimacy.<sup>61</sup> Because it thus results in lots of citizens being sent to jail for a long periods of time, the conventional deterrence strategy turns out not to be particularly cost-effective after all — not to mention morally problematic on a host of nonutilitarian grounds.

The contribution that social conditions of this sort make to street criminality — and the potentially perverse effect of the classical deterrence strategy on these conditions — can be systematized and refined by the reciprocity theory. The diverse psychological and social factors that predict crime suggest that reciprocity dynamics are at work within not just one but rather *three* interlocking collective action dynamics. The first consists in whatever mismatch exists between the interests of society in law-abiding behavior and the interests of individuals in committing crime. This is the *public order* collective action problem that occupies the attention of the conventional theory. The contribution that social influence makes to crime suggests that in this collective action setting as in others, many individuals behave like reciprocators: that is, they tend to respect the security of others in their persons and property in proportion to their perception that others are doing the same.<sup>62</sup>

The second collective action problem focuses on the collective good of *community self-policing*. Neighborhoods can do a lot to protect themselves from crime. Individuals can watch over one another’s residences. People can take an interest in the activities of one another’s children, alerting parents when they see neighborhood kids veering into trouble or

---

<sup>59</sup> See Tom R. Tyler, *Why People Obey the Law* (1990).

<sup>60</sup> See Kahan, *supra* note 15, at 359-60.

<sup>61</sup> See Jeffrey Fagan & Tracey L. Meares, *Punishment, Deterrence and Social Control: The Paradox of Punishment in Minority Communities* (Columbia Law School Public Law & Legal Theory Working Paper No. 10, Mar. 25, 2000).

<sup>62</sup> See Kahan, *supra* note 60.

even taking the effort to steer them out of it themselves.<sup>63</sup> Individuals can make their communities safer just by maintaining a conspicuous presence on its sidewalks and streets, especially at night.<sup>64</sup> It benefits the community collectively when everyone engages in these activities. Yet it remains in the interest of each individual to free-ride on the willingness of others to monitor, mentor, and simply hang out while attending exclusively to his or her own private business, especially where such activities can expose those who engage in them to risk or inconvenience.

The impact of *social organization* on crime suggests that reciprocity dynamics play a large role in determining how citizens respond to the community self-policing dilemma, too. Where they regularly encounter each other in voluntary associations — from churches to PTAs, from neighborhood improvement organizations to local chambers of commerce — citizens are much more likely to observe other individuals contributing to common endeavors and thereafter to reciprocate by doing the same. In atomized communities, in contrast, individuals are necessarily thrown back on their own devices; they are much less likely, in that circumstance, to see examples of public-spirited behavior and thus much less likely to fall into self-reinforcing patterns of common regard and concern.<sup>65</sup>

The third collective action problem hinges on the public good of *citizen-police cooperation*. The police obviously benefit when citizens cooperate with them by supplying them with information about crime.<sup>66</sup> Citizens benefit, too, when the police attend diligently to their needs and treat them with respect in daily encounters. Yet it will often be in the individual interest of citizens and police officers not to behave in these ways. When individuals report crimes, they expose themselves at a minimum to inconvenience, but also to risk of violent retaliation at the hands

---

<sup>63</sup> See generally Elijah Anderson, *Streetwise: Race, Class, and Change in an Urban Community* 3, 70-77 (1990) (discussing role of generalized youth supervision, and consequence of its deterioration, in containing crime in inner-city); Tracey L. Meares, *Social Organization and Drug Law Enforcement*, 35 *Am. Crim. L. Rev.* 191, 204, 207 (1998) (surveying empirical evidence).

<sup>64</sup> See Jane Jacobs, *The Death and Life of Great American Cities* 29-35 (1961)

<sup>65</sup> See Meares, *supra* note 63; Robert D. Putnam, *Bowling Alone: The Collapse and Revival of American Community* (2000).

<sup>66</sup> See generally Martín Sánchez Jankowski, *Islands in the Street: Gangs and American Urban Society* 193, 202-03 (1991) (arguing that cooperation between community and police is both necessary and sufficient to destroy viability of criminal gangs).



of those they reporting.<sup>67</sup> Where the law is perceived to be illegitimate, or enforcers arbitrary or biased, individuals who cooperate with the police are likely to experience personal guilt, or to be stigmatized by other members of the community.<sup>68</sup> For their part, the police might perceive that civilized and polite engagement with private citizens sometimes makes it harder for them to ferret out information necessary to solve crimes, or even exposes *them* to physical risk.<sup>69</sup> They might also prefer to avoid the risks and inconveniences associated with safeguarding private citizens from crime.

Reciprocity dynamics figure largely in a community's capacity to negotiate this collective action problem as well. Citizens are most disposed to cooperate with police when institutions enjoy a high level of *legitimacy*. Whether institutions are perceived as legitimate, it has been shown, is determined largely by whether citizens perceive they are being treated in a fair and respectful way by police and other decisionmakers.<sup>70</sup> In effect, citizens reciprocate respectful treatment with cooperation and obedience and disrespectful treatment with resistance — not only to the directives of individual decisionmakers but to the commands of the law more generally.<sup>71</sup> How compliant or resistant, deferential or defiant citizens are perceived to be no doubt influences the willingness of the police in turn to interact with them in a civil rather than a coercive fashion and otherwise respond attentively to their needs.<sup>72</sup>

The inefficacy of the conventional deterrence strategy is a consequence of the effects it has in promoting or inhibiting reciprocal cooperation within these three collective action settings. Considered in isolation, the effect of the conventional deterrence strategy on the *public-order* collective action problem is ambiguous. It's implausible to think that the

---

<sup>67</sup> See George Akerlof & Janet L. Yellen, *Gang Behavior, Law Enforcement and Community Values*, in *Values and Public Policy* 180 (Henry J. Aaron, Thomas E. Mann, & Timothy Taylor eds. 1994).

<sup>68</sup> See *id.* at 181-82; Anderson, *supra* note 63, at 190, 195-96, 205.

<sup>69</sup> See Tom R. Tyler, *Trust and Law Abidingness: A Proactive Model of Social Regulation*, 81 *B.U.L. Rev.* 361, 368-69 (2001).

<sup>70</sup> See *id.* at 367-68, 376-78, 385-86.

<sup>71</sup> See *id.* at 389.

<sup>72</sup> See Anderson, *supra* note 63, at 202-03; *cf.* Tyler, *supra* note 69, at 369, 384 (noting potential for displays of aggression to feed on each other in encounters between police and citizens).

threat of punishment has *no* restraining influence, particularly on individuals who for whatever reason are *not* restrained by the socially inculcated dispositions such as shame and guilt.<sup>73</sup> At the same time, as the effect of high-profile tax auditing campaigns suggests, it seems reasonable to infer that conspicuously severe penalties for street crimes might sometimes operate as a *cue* that criminality is in fact wide spread, an inference that through reciprocity dynamics would dilute the motivation of some individuals to respect the rights of others.

But even assuming that its effect on the public order dilemma is positive on the whole, the classical deterrence strategy clearly has a *negative* effect on the community-self-policing and the citizen-police-cooperation dilemmas. Public law-enforcement and community self-policing are, economically speaking, substitutes for one another. That is, the more a community has of one, the less it needs of the other in order to hold crime in check. Accordingly, as the state purports to assume a larger share of the deterrence burden through adoption of severe penalties, it actually *undermines* the incentive that individuals have to collaborate with each other to safeguard their communities from crime, at least to an extent.<sup>74</sup> As public enforcement suppresses community self-policing in this way, citizens have less occasion to observe one another making conspicuous contributions to the safeguarding of their own communities from crime. And having less exposure to monitoring, mentoring, creating a street presence and so forth, individuals, as reciprocators, become even less inclined to engage in such behavior themselves.<sup>75</sup> In effect, severe penalties, *crowd out* and *mask* the disposition of individuals to contribute to community self-policing, making it all the more necessary to employ severe penalties.

Severe penalties also discourage individuals from cooperating with the police. Such penalties increase the likelihood that the targets of reporting will retaliate. Indeed, if severe penalties are used to compensate for a low certainty of detection and conviction, most individuals will per-

---

<sup>73</sup> See generally Harold G. Grasmick & Donald E. Green, *Legal Punishment, Social Disapproval and Internalization as Inhibitors of Illegal Behavior*, 71 *Crim. L. & Criminology* 325 (1980).

<sup>74</sup> See generally Tomas J. Philipson & Richard A. Posner, *The Economic Epidemiology of Crime*, 39 *J. L. & Econ.* 405 (1996); Keith Hylton, *Optimal Law Enforcement and Victim Precaution*, 27 *Rand J. Econ.* 197 (1996); Omri Ben-Shahar & Alon Harel, *Blaming The Victim: Optimal Incentives for Private Precautions Against Crime*, 11 *J. L. Econ. & Org.* 434 (1995).

<sup>75</sup> See Anderson, *supra* note 63, at 57-58.

ceive that the likelihood of obtaining any benefit from reporting is largely futile anyway. In addition, particularly in minority communities, severe penalties help to construct the perception that the system is unjust. Accordingly, it is when the state penalizes criminal-wrongdoing severely that individuals are most likely to be inhibited from cooperating out of guilt or fear of being branded a collaborator. Confronted with an uncooperative citizenry, the police are likely to respond by engaging in heavy-handed enforcement, either to compensate for the dearth of private tips, to protect their own security, or simply to vent their frustration. This behavior by the police will in turn provoke citizens to be even less cooperative. Deprived of the benefits associated with community support — which turns out to be the most potent weapon for combating gangs<sup>76</sup> — the state will be forced to resort to even more severe penalties, thereby aggravating the citizen-police cooperation problem all the more.<sup>77</sup>

Ultimately, the negative effect of the classical deterrence strategy on the community-self-policing and citizen-police-cooperation dilemmas vitiates whatever positive effect the strategy might have had on the public-order dilemma. Convinced that those in the community will not do anything to stop crime, and resentful of a heavy-handed state, individuals are likely to respond by engaging in more law-breaking, which then feeds on itself as the spectacle of rampant criminality induces others to abandon whatever compunction they might have felt not to prey on their neighbors. The result is a self-sustaining high crime-rate equilibrium, fueled by distrust and various forms of negative reciprocity.

Is there a strategy for combating street crime that we should expect to work better from a reciprocity point of view? There is — *viz.*, the selective delegation of law enforcement and punishment functions to networks of private anti-crime associations.

Chicago has implemented a model form of this type of community policing. Under CAPS — the Chicago Alternative Policing Strategy — the Chicago Police Department divided the city's most crime-ridden neighborhoods into a collection of "advisory councils," which usually comprised no more than two or three city blocks. Each council was assigned a "beat officer," who was under strict instructions (at a time when the Mayor desperately feared a successful challenge from a minority candidate) to translate the council's grievances into an agenda of prob-

---

<sup>76</sup> See Jankowski, *supra* note 66, at 202-03.

<sup>77</sup> See Akerlof & Yellen, *supra* note 67, at 192-93, 195.

lems to be solved by policing strategies acceptable to community residents.<sup>78</sup>

The strategies that turned out to be the most acceptable involved the selective privatization of a variety of law-enforcement tasks. One of these was order-maintenance policing. In events dubbed “Operation Beat Feet,” “March for Peace,” and “Good Guys Loitering,” the advisory councils organized large numbers of law-abiding citizens to occupy the streets of disorderly neighborhoods. By establishing a “positive people presence,” these citizens transformed those neighborhoods into law-abiding ones during hours when they otherwise might have been expected to be the center of criminal activity.<sup>79</sup>

CAPS also privatized criminal investigations. At advisory council “beat meetings,” citizens frequently complained of sources of disorder that the police lacked the resources to investigate. When this happened, the citizens themselves were encouraged to gather the evidence necessary to obtain legal relief. Thus, on one occasion citizens facilitated the closure of a noisy tavern, which attracted disorderly patrons, by furnishing evidence of chronic health-code violations. On another, citizens contributed to the jailing of a slumlord, whose rundown tenement had become the site of drug-dealing and gang activity, by collecting evidence of “reckless disregard” for public safety.<sup>80</sup>

Finally, CAPS facilitated instances of private shaming. One of these involved a two-year picketing campaign, in which homeowners demonstrated outside the home of a slumlord who had allowed his properties to become the sites of deadly gang activity. The demonstrators “were fed up with the noise, crime, violence and general unrest that stemmed from the problem buildings . . . . They hoped they could make the building owner as uncomfortable in his home as he was making them in theirs.”<sup>81</sup>

This form of highly participatory and decentralized law enforcement proved to be as successful as it was unorthodox. Examining crime and opinion data, criminologists Wesley Skogan and Susan Hartnett have concluded that in the neighborhoods in which CAPS operated trust in the po-

---

<sup>78</sup> See generally Wesley G. Skogan & Susan M. Hartnett, *Community Policing Chicago Style* (1997).

<sup>79</sup> See *id.* at 174-75, 225.

<sup>80</sup> See *id.* at 166-67, 175-76.

<sup>81</sup> *Id.* at 177-78.

lice grew significantly, as did trust among neighbors. All forms of street crime — from drug distribution to robbery to homicide — dropped.<sup>82</sup>

The behavioral mechanisms at work can again be explained in reciprocity terms. In effect, CAPS, promoted trust, and hence reciprocal cooperation, within each of the three collective action settings that construct the problem of street crime. To begin, CAPS had a positive effect on the community-self-policing dilemma. Whereas traditional policing strategies risk displacing community self-policing, CAPS assigned certain, highly conspicuous elements of law-enforcement to community residents themselves. As they observed their neighbors attending and speaking up at council meetings, and thereafter participating in order-maintenance demonstrations, public shamings, and the like, citizens learned that in fact their neighbors *were* willing to take an active role in safeguarding their community from crime. Those who formed this impression could thereafter have been expected to reciprocate, either by participating in CAPS initiatives or by entering into less formal arrangements to watch out for one another's interests.

The CAPS approach to community policing also helped to promote positive reciprocity within citizen-police-cooperation setting. Citizens long accustomed to seeing the police as simultaneously indifferent to their needs and disrespectful of their rights now were exposed to highly responsive and solicitous officers. Unsurprisingly, citizens grew more trustful and thus more willing to cooperate with the police. In addition, CAPS made it easier to cooperate with the police by negating social meanings that can make such behavior an occasion for guilt or ostracism: those who took part in CAPS were not likely to view themselves or to be viewed by others as turning their fellow citizens over to an alien or occupying force; rather they were participating in forms of self-governance visibly supported by other members of the community. The police, too, no doubt reciprocated the greater willingness of citizens to cooperate with them by treating citizens more respectfully in return, thereby generating an even greater willingness among citizens to cooperate with the police.

Because it had these effects on the community-self-policing and citizen-police-cooperation problems, CAPS likely had a positive effect on the public-order dilemma as well. In a climate in which they trusted each other and the state more, individuals are more likely to obey the law. And through reciprocity dynamics, such obedience feeds on itself.

The reciprocity theory explains why we should expect selective privatization to result in a self-sustaining high-cooperation, low-crime equilib-

---

<sup>82</sup> *See id.*

rium. And it implies that this equilibrium is likely to be a stable and lasting one.

## 6. Other Applications

The reciprocity theory has implications for a broad range of policy problems in addition to tax collection and the siting of noxious facilities. It's possible to sketch several in broad outline.

### *Fraud and Corruption*

Like the disposition of individuals to engage in tax evasion, the disposition of individuals to engage in fraud or corruption appears to depend on whether they think other individuals are engaged in such behavior.<sup>83</sup> This implies that high-profile campaigns to crackdown on such behavior, like high-profile crackdowns on tax evasion, can backfire.<sup>84</sup> Indeed, when government invests more to deter fraud, individuals have less incentive to invest in credibly signaling to others that they are trustworthy and honest, and hence reliable as trade partners. Because individuals reciprocate honesty with honesty, the suppression of individuals efforts to display honesty to others will predictably reduce the disposition of individuals to behave honestly, thus making penalties for dishonesty less effective. A better policy, again, is to make citizens aware that those around them are basically honest.

Or at least that is the best policy where individuals are in fact generally honest. In a condition of pervasive distrust — such as that which obtains in many former Eastern block nations — strong penalties for fraud and dishonesty may be the only thing that works. Moreover, in such a climate, penalties for dishonesty may in fact promote rather than undermine trust. Individuals who resent fraud and corruption are likely to interpret the advent of credible penalties as evidence that others around them now feel the same way and are prepared to do something about it. Some of those individuals will be moved to reciprocate by behaving more honestly themselves, inducing still others to do the same, and so forth and so

---

<sup>83</sup> See Jon Elster, *The Cement of Society: A Study of Social Order* 278-70 (1989); Peter H. Huang & Ho-Mou Wu, *More Order Without Law: A Theory of Social Norms and Organizational Cultures*, 10 J.L. Econ. Org. 390 (1994)..

<sup>84</sup> Elster, *supra* note 83, at 270.

on, until a new condition of self-reinforcing cooperation is reached — at which point maintenance of high penalties may be less necessary.<sup>85</sup>

### *Information and Technology*

Ideas are understood to be a classic public good. We all benefit from useful inventions, engaging works of literature, effective medicines for disease, and the like. But why should any *one* of us endure the cost associated with producing them when we can freely avail ourselves of the inventive labors borne by others? The conventional theory again resorts to incentives, here in the form of intellectual property rights that permit inventors to exclude others from use of their ideas absent the payment of a fee.<sup>86</sup>

But reciprocity once more complicates the picture. A growing body of work has documented that within certain fields — including basic scientific research and many types of computer software development — individuals will reciprocate spontaneous contributions to a collaborative inventive enterprise with like contributions of their own, generating innovations that rival and often surpass the quality of those achieved through proprietary modes of production.<sup>87</sup> Where this is so, the deadweight losses and administrative costs inevitably associated with intellectual property rights needn't be endured to secure the public benefits of invention. Indeed, university scientists, computer hackers, and other reciprocal producers tend to suspend the free exchange of ideas once they come to suspect that those with whom they are collaborating are intent on appropriating the commercial value of those innovations for themselves.<sup>88</sup> An intellectual property regime that is insensitive to the contribution that reciprocity norms make to invention can thus stifle rather than stimulate innovation.

---

<sup>85</sup> See generally Susan Rose-Ackerman, *Corruption and Government: Causes, Consequences, and Reform* (1999).

<sup>86</sup> See generally Kenneth J. Arrow, *Economic Welfare and the Allocation of Resources for Invention*, in *The Rate and Direction of Inventive Activity: Economic and Social Factors* 609 (1962); Harold Demsetz, *The Private Production of Public Goods*, 13 *J.L. & Econ.* 293 (1970).

<sup>87</sup> See, e.g., Yochai Benkler, *Coase's Penguin, or, Linux and the Nature of the Firm*, Yale L.J. (forthcoming 2003); Arti Kaur Rai, *Regulating Scientific Research: Intellectual Property Rights and the Norms of Science*, 94 *Nw. U.L. Rev.* 77 (1999).

<sup>88</sup> See Benkler, *supra* note 87; Rai, *supra* note 87.

*Democracy*

The application of the conventional model of collective action to democratic politics yields public choice theory. According to that theory, citizens, because they are self-interested wealth maximizers, will forego public spirited deliberation and instead organize themselves into interest groups for the purpose of extracting rents.<sup>89</sup> To combat this dynamic, policy analysts have proposed a wide variety of structural devices — from campaign finance laws<sup>90</sup> to term limits<sup>91</sup> to line item vetoes<sup>92</sup> to budget process reforms<sup>93</sup> — all of which seek to raise the cost or reduce the benefits of organizing into special-interest pressure groups.

The reciprocity model suggests a different analysis. As a positive matter, it points to a substantial body of empirical research suggesting that the behavior of elected representatives is limited by informal norms that discourage unconstrained efforts to redirect public resources toward one's own constituencies.<sup>94</sup> Thus, reciprocity dynamics already make at least some contribution to containing special-interest politics.

As a prescriptive matter, the reciprocity model warns us not to assume that structural reforms will invariably reinforce reciprocity norms in this setting. Policies designed to counteract public choice pressures do more than change political actors' incentives to engage in rent-seeking; they also broadcast to citizens and their representatives that rent-seeking is the behavior we *expect* political actors to engage in whenever it is in their interest to do so. Because individuals are reciprocators, they are likely to respond to this message by displaying even *less* restraint in the pursuit of their material interests in democratic political life. Thus, reforms

---

<sup>89</sup> Olson, note 1 above, is again the foundational work. See also James M. Buchanan & Gordon Tullock, *The Calculus of Consent: Logical Foundations of Constitutional Democracy* (1962).

<sup>90</sup> See, e.g., Ian Ayres & Jeremy Bulow, *The Donation Booth: Mandating Donor Anonymity to Disrupt the Market for Political Influence*, 50 *Stan. L. Rev.* 837 (1998).

<sup>91</sup> See generally Elizabeth Garrett, *Term Limitations and the Myth of the Citizen-Legislator*, 81 *Cornell L. Rev.* 623 (1996) (critiquing use of term limits to counteract public choice dynamics).

<sup>92</sup> See Elizabeth Garrett, *Accountability and Restraint: The Federal Budget Process and the Line Item Veto Act*, 20 *Cardozo L. Rev.* 871 (1999).

<sup>93</sup> See Elizabeth Garrett, *Rethinking the Structures of Decisionmaking in the Federal Budget Process*, 35 *Harv. J. Leg.* 1113 (1998).

<sup>94</sup> See generally Donald P. Green & Ian Shapiro, *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science* (1994).



aimed at reducing incentives to behave in a self-interested fashion might well dissipate reciprocity-based norms that now hold such behavior at least partially in check, and thereby *increase* special-interest rent-seeking on net. The reciprocity model thus underscores the anxiety that too readily *accepting* the public choice picture can *make* it the reality of our political life.<sup>95</sup>

At the same time, however, the reciprocity model underscores how reforms that reflect different assumptions might stimulate public spiritedness. For example, scholars have proposed that the state award citizens two types of monetary grants: “stakes” that they can use as they see fit upon adulthood, and “patriot dollars” that they can contribute to the political campaigns of their choice.<sup>96</sup> The first of these expresses a societal commitment to assuring individuals a fair chance to realize their life plans, the second its commitment to assuring them a fair chance to influence the political process, irrespective of their personal wealth. It’s plausible to believe that many citizens will reciprocate the good will embodied in these schemes by contributing more readily to the well-being of society and by refraining from purely self-seeking political behavior. And when they observe public-spirited behavior of this sort, still more citizens will be moved to behave in the same way. These proposals, then, are another example of how appropriately expressive law — even in the form of cash subsidies — can be expected to accentuate reciprocal cooperation.

#### *Good Samaritanism*

Breaking with the traditional Anglo-American position, several states have in recent years enacted laws that oblige individuals to assist strangers in need when they can do so without risk to themselves. Such laws are intended to counter the supposed growing indifference of Americans — particularly urban-dwelling ones — toward the well-being of strangers.<sup>97</sup>

But the reciprocity theory warns that such laws may do more to construct than to remedy such indifference. Some individuals will see the apparent necessity of a penalty for nonassistance as confirmation that most citizens don’t genuinely care about strangers’ well-being; those indi-

---

<sup>95</sup> See Jerry Mashaw, *Greed, Chaos, and Governance: Using Public Choice to Improve Public Law* (1997).

<sup>96</sup> Bruce Ackerman & Anne Alstott, *The Stakeholder Society* (1999); Bruce Ackerman & Ian Ayres, *Voting with Dollars* (forthcoming, Yale Univ. Press 2002).

<sup>97</sup> See Daniel B. Yeager, *A Radical Community Of Aid: a Rejoinder to Opponents of Affirmative Duties to Help Strangers*, 71 Wash. U.L.Q. 1 (1993).

viduals, the reciprocity model predicts, will respond by showing less concern themselves. Financial incentives to assist are also likely to obscure *morally* motivated acts of assistance, thereby diluting a signal of good intentions that would otherwise have moved individuals to reciprocate in kind.

Substantial experimental evidence suggests that it simply is not the case that Americans are disinclined to render assistance to strangers in need.<sup>98</sup> The way to strengthen citizens' resolve to render such assistance, the reciprocity theory implies, is to correct the misperception that others lack such resolve, a goal that can be achieved through public commendations of individuals who engage in heroic behavior.

## 7. Conclusion

The main — indeed, only — selling point of the conventional theory of collective action is its asserted behavioral realism. Individuals, it tells us, are inherently self-seeking. Accordingly, we can't count on them voluntarily to subordinate their material interests to the good of society; rather we must alternately bribe and threaten them through a costly regulatory apparatus, the maintenance of which not only depletes our common resources but itself creates myriad opportunities for advantage-seeking by self-interested individuals and groups. It is hard to imagine a less inspiring account of our motives and our prospects. But if the ugly picture the conventional theory paints is *right*, then we'd be fools to avert our eyes from it.

It turns out, however, that the conventional theory *isn't* right. Individuals in collective action settings might not behave like saints, but they don't behave like fiends either. They *can* be counted on to contribute to collective goods, the emerging literature on reciprocity shows, *so long as they perceive that others are inclined to do the same*. Bribes and threats are not nearly so necessary as the conventional theory would have us believe; the law can instead enlist our cooperation by furnishing us with grounds to *trust* one another to contribute our fair share to society's needs. Indeed, when the law relies only on bribes and threats, it breeds the impression that citizens can't trust one another to contribute to collective goods voluntarily, thereby undermining their motivation to reciprocate one another's public spiritedness. Whatever truth there is in the

---

<sup>98</sup> See Bibb Latane & John M. Darley, *The Unresponsive Bystander: Why Doesn't He Help?* (1970) (reporting experimental results showing that failure to intervene is attributable to errors in perception especially likely to occur in group settings).

conventional theory is an artifact of the common acceptance of that theory's bleak assumptions.

So we should now reject them. To replace the conventional theory of collective action, we should construct a new and more appealing one founded on our nature as reciprocators. The logic of reciprocity not only reflects a more realistic understanding of individual emotional and moral commitments. It makes the hope that citizens will be morally and emotionally committed to contribute to the common good more realistic.