VIRGINIA LAW REVIEW

VOLUME 86

NOVEMBER 2000

NUMBER 8

DO GOOD LAWS MAKE GOOD CITIZENS? AN ECONOMIC ANALYSIS OF INTERNALIZED NORMS

*Robert Cooter**

INTRODUCTION

DOES an effective state require citizens who are virtuous or merely rational? Gerard Lynch recently articulated the answer of common sense and pre-modern scholarship:¹

What society wants from its members, in any case, is not an intelligent calculation of the costs and benefits of abiding by its basic norms, but more or less unthinking obedience to them. To the extent people are specifically comparing the costs and benefits of breaking criminal laws, the battle is already lost; many of them must conclude, in particular situations, that the calculus favors law-breaking For society to function, most people have to obey the law for reasons of conscience and conviction, and not out of fear of punishment.²

Against this view stands the economic analysis of law. Some economists proclaim themselves to be moral skeptics. Almost all economists, however, practice moral skepticism by exploring effi-

* Herman Selvin Professor of Law and Director of John Olin Program in Law and Economics, University of California at Berkeley, rdc@law.berkeley.edu. For useful comments, I would like to thank participants in the conference, "Legal Construction of Norms," sponsored by the University of Virginia John M. Olin Program in Law and Economics, and also the joint seminar for the political science department at the University of California at San Diego and the law school of the University of San Diego.

¹ In a private communication with me, legal historian James Gordley asserts that few pre-modern scholars questioned the belief that the health of the state reflects the virtue of its citizens. As a possible exception, he offers Machiavelli.

² Gerard E. Lynch, *The Role of Criminal Law in Policing Corporate Misconduct*, 60 *Law & Contemp. Probs.*, 23, 46 (1997).

cient institutional arrangements for rationally self-interested actors.

The success of the economic analysis of law demonstrates the power of skeptical models. Even so, skepticism fails to explain significant activities of people, as illustrated by four examples. First, in laboratory games conducted by economists and social psychologists, people persistently cooperate more than predicted by rational self-interest.³ For example, people share payoffs with others when doing so is unnecessary. Second, in general elections, the probability that a particular vote proves decisive is so small that purely self-interested citizens would not bother to vote at current rates.⁴ Whereas journalists chide citizens for low rates of participation in elections, economists find participation rates inexplicably high. Third, lawyers in the United States often abandon more lucrative practices to become federal judges, presumably because they enjoy participating in public life and shaping the law to their political vision. Fourth, in most countries the punishment for tax evasion, discounted by the probability of prosecution and conviction, is

³ See Max H. Bazerman & Margaret A. Neale, *The Role of Fairness Considerations and Relationships in a Judgmental Perspective of Negotiation*, in *Barriers to Conflict Resolution* 86 (Kenneth Arrow et al. eds., 1995); Karen Cook & Karen Hegvedt, *Empirical Evidence of the Sense of Justice*, in *The Sense of Justice: Biological Foundations of Law* 187, 197–200 (Roger D. Master & Margaret Gruter eds., 1992); Elizabeth Hoffman et al., *Behavioral Foundations of Reciprocity: Experimental Economics and Evolutionary Psychology*, 36 *Econ. Inquiry* 335, 347 (1998); Elizabeth Hoffman et al., *Preferences, Property Rights and Anonymity in Bargaining Games*, 7 *Games & Econ. Behav.* 346, 371–72 (1994); Elizabeth Hoffman & Matthew L. Spitzer, *Entitlements, Rights, and Fairness: An Experimental Examination of Subjects' Concepts of Distributive Justice*, 14 *J. Legal Stud.* 259, 259–60 (1985); Matthew Rabin, *Incorporating Fairness into Game Theory and Economics*, 83 *Am. Econ. Rev.* 1281, 1283 (1993); Paul Burrows & Graham Loomes, *The Impact of Fairness on Bargaining Behavior*, presented at European Association of Labour Economists annual meeting, Rome (Sept. 20, 1990).

⁴ According to one calculation, the probability of casting a decisive vote in a typical U.S. general election approximately equals 10^{-8} . See Richard L. Hasen, *Voting Without Law?*, 144 *U. Pa. L. Rev.* 2135, 2141 n.25 (1996). Using a different method of calculation, Romer concludes that the probability of a tie in a U.S. presidential election in which 50 million people vote is approximately 10^{-4} . See Paul M. Romer, *Preferences, Promises, and the Politics of Entitlement in Individual and Social Responsibility: Child Care, Education, Medical Care, and Long-Term Care in America* 195, 200 (Victor R. Fuchs ed., 1996). For a comparison of self-interested and civic-minded theories of voter participation, see Robert D. Cooter, *The Strategic Constitution* 19–23 (2000). A more careful strategic theory of voting developed by Timothy Feddersen and Wolfgang Pesendorfer provides a clever but ultimately unconvincing solution to this puzzle. See Timothy J. Feddersen & Wolfgang Pesendorfer, *Abstention in Elections with Asymmetric Information and Diverse Preferences*, 93 *Am. Pol. Sci. Rev.* 381 (1999).

small relative to the gain. Whereas economic models of self-interest predict low rates of tax compliance, some countries, like the United States and Switzerland, enjoy unusually high rates of tax compliance.⁵

To encompass these facts, economists can postulate "tastes" for fairness, voting, judging, and tax compliance. Or economists can postulate general tastes like "self-expression" and "good citizenship" that encompass these particular tastes. Postulating a taste for morality raises the question, "What is the difference between an unselfish desire to treat others fairly and a selfish desire to satisfy a taste for treating others fairly?" While philosophers puzzle over whether seemingly good people are selfishly satisfying a taste for morality, this question is unimportant for social science. Social scientists should describe the values internalized by people, predict the effects of internalized values on society, and explain why some people internalize values that others do not internalize. In brief, social scientists should chart the distribution, effects, and causes of internalized values.

I will extend economics to sketch such a theory. My focus will be on social norms that regulate civic acts. Examples include participating in government, helping officials enforce laws, encouraging honest administration, cleaning the local park, helping the needy, following the rules, and treating others fairly.

When building in Rome, excavation inevitably exposes the detritus of past civilizations. Similarly, a long, contentious history attaches to phrases like "social norms" and provokes confusion over definitions. In my search for causes, I leave argument over definitions to others, but I will use an old tradition in Anglo-American jurisprudence to provide the minimum meaning needed to motivate my model. This tradition, called the "imperative theory of law," asserts that a law is an obligation backed by a state sanction.⁶ According to this tradition, the law consists of statements

⁵ Eric Posner explains the problem and attempts a solution. See Eric Posner, *Law and Social Norms: The Case of Tax Compliance*, 86 Va. L. Rev. 1781 (2000).

⁶ In its early form in English jurisprudence, the positive theory asserts that law is the command of a sovereign, where a sovereign is someone whom others obey and who does not obey anyone. Generalizing, the "sovereign" becomes the process for making laws stipulated in higher-level laws. A law created by following the correct procedures has the right "pedigree." See Ronald Dworkin, *Taking Rights Seriously* 17 (1977). In the modern formulation, a law regulating behavior is an effective

such as, "You ought not to drive over 70 miles per hour, or else you must pay a fine of \$500," or "You ought to take reasonable care towards your neighbor, or else you must pay compensation for the harm that you cause." Following this tradition, a norm can be defined as an obligation backed by a *nonlegal* sanction. Sanctions such as criticizing, blaming, refusing to deal or shunning are nonlegal insofar as the people who impose them are not state officials.

Laws also differ from some social norms in another respect besides the sanction. According to the U.S. Constitution, a bill becomes law when it receives a majority vote in both houses of Congress and the President signs it.⁷ In general, legal systems have rules for making, amending, or extinguishing rules. Similarly, the charter of a corporation may empower its board of directors to impose rules on employees, and the bylaws of a church may empower its deacons to impose rules on church members. Like the state, many private organizations have charters that specify how to make rules. The least formal social norms, however, lack rules for making rules. For example, no definite process exists to create, amend, or extinguish a rule of etiquette or a principle of morality.

This fact contributes to disagreements about the existence and content of morality and etiquette, especially in times of social change. To illustrate, are Americans still obligated to hold the fork in the right hand when eating, or is the European practice of holding the fork in the left hand now acceptable in America? This question apparently has no definitive answer. Disagreement over what norms exist, however, is modest compared to disagreement over what norms ought to exist. To illustrate, some conservative religious groups teach women to obey their husbands, whereas feminists cannot tolerate the subordination of women to men. Conventional morality asserts that people ought to recognize obligations that critical morality denies.

Instead of discussing what norms exist or ought to exist, I focus on the interaction between actual norms and laws. To understand the interaction between law and norms, social scientists should describe the distribution and effects of internalized values. In consumer theory, the amount that a person is willing to pay for a

obligation with the right pedigree. See H.L.A. Hart, *The Concept of Law* 107-14 (1961); Joseph Raz, *The Concept of a Legal System: An Introduction to the Theory of a Legal System*, 147-56 (2d ed. 1980).

⁷ See U.S. Const. art. I, § 7, cl.2.

good measures the strength of his preference for it. Similarly, I will measure the extent to which a person internalizes a social norm by the amount that he will pay to conform to it. Next I will show how the interaction between willingness-to-pay and the actual cost of conforming to a norm determines the equilibrium level of obedience. Finally, I will show that when some citizens internalize respect for law, pronouncement of a new law can have an expressive effect that causes behavior to jump to a new equilibrium. Given appropriate internalization, legal expression changes behavior dramatically with little state expenditure on coercion.

In the Paper's second half, I will extend economics to explain the causes of internalization. In many cooperative ventures, people prefer partners committed to moral values. Rationally self-interested people, consequently, internalize moral values in order to increase their opportunities for cooperation with others. A Pareto self-improvement is a change in values that increases opportunities enough so that the outcome is better relative to the actor's original and final preferences. People make Pareto self-improvements in response to enduring relationships, especially those involving intimacy. State officials who have remote relationships with citizens lack the information needed to detect internalized values. Instead of building internalization, the state mostly relies on citizens' families, friends, and colleagues to encourage civic virtue.

I. WILLINGNESS-TO-PAY

Performing civic acts often costs money, time, effort, opportunities, inconvenience, or risk. The vertical axis in Figure 1 represents the amount a person would be willing to pay to do a particular civic act, and the horizontal axis represents the proportion of citizens willing to pay the price. According to the graph, roughly 80% of the citizens will pay something to do their civic duty, whereas roughly 20% will pay nothing. Thus roughly 80% of the population internalizes their civic duty (although not to the same extent), and roughly 20% of the population externalizes their civic duty. Imagine a state in which 80% of the citizens will expend some time and effort in order to vote, but not everyone is willing to expend the same amount, and 20% of the citizens are unwilling to expend anything. Thus the willingness-to-pay curve describes the distribution of tastes for voting.

Figure 1: Willingness to Pay

The theory of consumer demand elucidates Figure 1. For some goods such as pizzas or gasoline, a fall in price causes some people to buy more of the good. For other goods, such as houses and dentures, most people buy only one good at most. For the latter, a fall in price causes more people to buy the good. Figure 1 resembles a demand curve for goods of the second type.

Another distinction concerns the difference between derived demand and final demand. The demand of producers for goods used in production derives from the demand of consumers for goods used in consumption. For example, the demand for seed corn, chromium, and sulfuric acid by producers derives from the demand by consumers for such goods as flour, automobiles, and newspapers. In general, the preferences of consumers determine final demand, and final demand determines derived demand.

To understand the importance of this difference, assume that someone wants a car for personal enjoyment and work. If the personal enjoyment is worth \$10,000 and the use for work is worth \$5,000, then the person is willing to pay \$15,000 for the car. An

economist can observe that the person is willing to pay \$15,000 for the car more easily than observing the breakdown into \$10,000 for personal use and \$5,000 for work. Discerning the difference between derived and final demand, however, is necessary for predictions. If I change jobs, my willingness to pay to use the car for work may fall from \$5,000 to \$0, but I may still be willing to pay \$10,000 to use the car for personal enjoyment. In this example, the economist must distinguish derived demand from final demand in order to predict the consequences of changing jobs on the demand for cars. Besides being important for predictions, discerning the difference between derived and final demand is necessary to evaluate human welfare, especially to criticize a Soviet-style economy that apparently produced steel for steel's sake.

The difference in causes between derived demand and final demand rests on the difference between instrumental and intrinsic values. Relatively instrumental values motivate derived demand, and relatively intrinsic values motivate final demand. These values are "relative" because so many goods have both instrumental and intrinsic values. Economists describe relatively intrinsic values as "tastes" or "preferences." The preferences of people determine what goods they want to buy, while technology and institutions determine costs.

In my analogy between norms and markets, final demand for a commodity corresponds to the intrinsic value of obeying a social norm. A person who intrinsically values obeying a social norm will pay something to obey the norm for its own sake, independent of any resulting advantage or disadvantage. In the language of economics, intrinsic value implies a "taste" for obeying the norm. Tastes are internal to people. The curve in Figure 1 depicts the distribution of internalized values among citizens.

While the values depicted in Figure 1 are relatively intrinsic, civic acts can also have instrumental value. The instrumental value of civic acts often depends on the advantage gained from having the reputation of being a good citizen. A businessman may participate in civic projects in the hope that a good reputation prompts leniency from state regulators. As with owning a car, the intrinsic and instrumental values of obeying a social norm often get confused, even in the actor's mind. Distinguishing them, however, is essential to accurate predictions. For the businessman, increasing

his anonymity undermines the importance of reputation, which diminishes the instrumental value of civic acts, but does not diminish their intrinsic value.

Figure 1 represents the intrinsic value of civic acts. Following this interpretation, Figure 1 indicates that roughly 80% of the population has a taste for obeying the norm and roughly 20% of the population does not have this taste. To illustrate concretely, 80% of the population may be willing to spend time voting. These people presumably believe that they ought to vote. However, 20% of the people are unwilling to spend time voting. These people only vote if doing so has instrumental value, such as winning praise from neighbors or securing advantages from politicians.

Whereas Figure 1 depicts a situation where most people intrinsically value obeying a norm, people could value obeying a norm entirely for instrumental reasons. For example, unlike Japanese lawyers, most American lawyers probably view hostile takeovers as morally neutral. Insofar as participating in hostile takeovers makes more money for them, they will participate. Conversely, if their participation loses money for them because firms refuse to deal with them, they will not participate. To represent people who place little intrinsic value on the norm, the willingness-to-pay curve in Figure 1 must shift down.

By restricting the interpretation of Figure 1 to final demand, intrinsic and instrumental value are graphically separated: The willingness-to-pay curve represents intrinsic value, and the cost curve, my next topic, represents all instrumental values.

II. COST

Having graphed willingness-to-pay to obey the norm, I now turn to the cost of obeying it. Obeying a norm often imposes direct costs in money, inconvenience, effort, risk, or lost opportunity. In addition, obeying a norm often conveys the benefits of a good reputation, such as praise, esteem, promotion, and preferential dealings. Finally, obeying a norm conveys the benefit of avoiding a social sanction. Combining these elements yields this formula for the net cost of obeying a social norm:

$$\text{net cost} = \text{direct cost} - \text{reputational benefit} - \text{avoided sanction}.$$

In general, net costs can be positive or negative, depending on the relative size of these three elements. With negative net costs,

everyone obeys the norm. I will focus on the usual case where net costs are positive and some people violate the norm.

Figure 2: Cost

Like industrial production, social norms can exhibit increasing returns to scale because of fixed costs and network effects.⁸ Social norms, however, can also exhibit increasing returns to scale for more distinctive reasons. People are notoriously susceptible to group pressures, variously described as conformity, herd effects, or social solidarity.⁹ With group pressures, an increase in an act's popularity lowers its cost. Imposing a nonlegal sanction on someone often involves a risk of retaliation, which decreases as more

⁸ To illustrate, the cost of digging a hole falls dramatically when enough holes must be dug to justify changing from a shovel to mechanized equipment. Mechanized equipment has fixed costs that are prominent in communication networks. Network effects in law are discussed in Marcel Kahan & Michael Klausner, *Standardization and Innovation in Corporate Contracting* (or "The Economics of Boilerplate"), 83 Va. L. Rev. 713 (1997); Michael Klausner, *Corporations, Corporate Law, and Networks of Contracts*, 81 Va. L. Rev. 757 (1995).

⁹ For an especially interesting attempt to understand these effects through economics, see Geoffrey Brennan & Philip Pettit, *The Peculiar Economy of Esteem*, Conference on Norms and Incentives, Australian National Research School of Social Sciences (July 20–21, 1998).

people obey the norm. The risk of a nonlegal sanction often increases as more people obey a norm, thus lowering the relative cost of conforming to the norm. The increased aggression of nonsmokers toward smokers as fewer people smoke in airports demonstrates this phenomenon.

Figure 3 combines the two curves. To reduce the number of graphs, I have drawn shapes for the curves in Figure 3 that capture the most interesting possibilities. Where the two curves intersect, the cost of doing the civic act equals the price people are willing to pay, so the system is in equilibrium. Stable equilibria occur in Figure 3 at 0% and roughly 50%, whereas an unstable equilibrium occurs at roughly 20%. Social interactions in Figure 3 have multiple equilibria whose analysis will yield some surprising results.

Figure 3: Equilibria

Consider the direction of movement when the system is out of equilibrium. Where the willingness-to-pay curve is above the expected cost curve, more people are willing to do the act than required to sustain the current cost of doing it, so the cost decreases. In the interval between 20% and 50%, the proportion of

people acting increases as indicated by the arrow in Figure 3. Conversely, where the cost curve is above the willingness-to-pay curve, fewer people are willing to do the act than required to sustain the current cost of doing it, so the cost is increasing. In the interval between 50% and 100%, the proportion of people acting decreases as indicated by the arrow. Similarly, in the interval between 0% and 20%, the proportion of people acting decreases as indicated by the arrow.

Thus, starting from any point below 20%, the system tends to move to 0%, and starting from any point above 20%, the system tends to move to 50%. Zero and 50% are the stable equilibria of the system depicted in Figure 3, and 20% is a tipping point. In Part III, I will use these facts to distinguish two distinct ways that internalization influences behavior. First, however, I will use the graph to explain the consequences of the absence of internalization.

Without internalization, the willingness-to-pay curve shifts down until it rests on the horizontal axis, as depicted in

Figure 4. In these circumstances, no one will obey the norm unless the *net* costs are negative. In other words, no one will obey the norm unless the instrumental benefit exceeds the direct cost. To depict a situation where someone might obey the norm, I also need to shift the net cost curve down so that it assumes negative values. In Figure 4, the negative values are assumed after roughly 40% or more of the population obeys the norm.

Figure 4: No Internalization Causes Instability

Notice that without internalization, the only stable equilibria in Figure 4 are the extremes where 0% obey the norm or 100% obey the norm. Without internalization no one will bear the cost of being different, so everyone does the same thing. This fact can cause jumps in behavior from one extreme to the other. For example, lawyers presumably have no moral commitment to participating or not participating in hostile takeovers, and are willing to do whatever is more profitable. Without internalization, the behavior of lawyers could swing from one extreme to the other. In fact, the situation has changed abruptly in the United States from no lawyers willing to participate in hostile takeover to all lawyers willing to participate.

III. TWO EFFECTS OF INTERNALIZATION

Return to Figure 3 and consider the effects of more internalization, which shifts up the willingness-to-pay curve as depicted in Figure 5. As a consequence of this shift, the tipping point at 20% moves down to approximately 15%. Similarly, the stable internal equilibrium at 50% moves up to approximately 65%. Thus the sys-

tem will now converge to 65% from any point above 15%, whereas before the system would converge to 50% from any point above 20%.

More internalization thus has one of two possible effects on conforming to norms in the long run. If the system were originally tending towards equilibrium at 50% conformity, then more internalization causes the system to tend toward 65% conformity. Alternatively, if the system were stuck at 0% conformity, then more internalization decreases the size of the shock from 20% to 15%, which causes the system to jump to a high level (65% and 50%, respectively). In general, more internalization causes either a small increase in conformity for certain, or a large jump in conformity with higher probability.

Figure 5: Internalization

What determines whether the system settles into equilibrium at 0% or 65%? If everyone believes that less than 15% of the citizens will do the act in question, then their belief will prove correct and the system will converge to the equilibrium at 0%. Conversely, if everyone believes that more than 15% of the citizens will do the act in question, then their belief will prove correct and the system

will converge to the equilibrium at 65%. So the system has characteristics of a self-fulfilling prophecy.

Given this fact, the law might play a crucial role in determining the outcome by influencing people's beliefs about what others will do. Recall that the figures represent a civic act. The state presumably wants citizens to do their civic duty. If the state is careful about its pronouncements, so that most citizens believe them, then the state might cause the system to converge to an equilibrium at 65% merely by making the appropriate pronouncement. In general, a credible state can influence citizen choice among multiple equilibria by pronouncing the law. Enacting "pooper-scooper" ordinances in Berkeley apparently caused many owners to clean up after their dogs, even though legal enforcement remains lax or absent. Similarly, posting "no smoking" signs in U.S. airports has improved the quality of indoor air with little or no legal enforcement. These ideas, which I have developed elsewhere,¹⁰ could revive older theories of expressive law by providing them with an analytical basis.¹¹

The graphs reveal the fundamental error of asserting that internalized values do not influence aggregate behavior. As depicted in Figure 3, willingness to pay and costs *jointly* determine the aggregate level of civic acts. In contrast, pure skepticism asserts that the cost curve solely determines the aggregate level of civic acts.¹² Marx's labor theory of value is the most famous example of this mistake in economics.¹³ Pure skeptics mistakenly suppose that the

¹⁰ See Robert Cooter, *Expressive Law and Economics*, 27 J. Legal Stud., 585, 593–96 (1998) [hereinafter Cooter, *Expressive Law*].

¹¹ See Joel Feinberg, *Doing & Deserving: Essays in the Theory of Responsibility* 95–118 (1970); Mathew D. Adler, *Expressive Theories of Law: A Skeptical Overview*, 148 U. Pa. L. Rev. 1363 (2000); Robert Cooter, *Self-Control and Self-Improvement for the "Bad Man" of Holmes*, 78 B.U. L. Rev. 903, 905–14 (1998) [hereinafter Cooter, *"Bad Man" of Holmes*]; Jean Hampton, *An Expressive Theory of Retribution*, in *Retributivism and its Critics* 1 (Wesley Cragg ed., 1992); Dan M. Kahan, *Social Influence, Social Meaning, and Deterrence*, 83 Va. L. Rev. 349, 350–51 (1997); Dan M. Kahan, *Social Meaning and the Economic Analysis of Crime*, 27 J. Legal Stud. 609 (1998); Richard H. Pildes, *Why Rights Are Not Trumps: Social Meanings, Expressive Harms, and Constitutionalism*, 27 J. Legal Stud. 725 (1998); Cass R. Sunstein, *On the Expressive Function of Law*, 144 U. Pa. L. Rev. 2021, 2021–25 (1996).

¹² This conclusion requires a vertical demand curve (perfectly inelastic demand). There is no more reason to think that aggregate levels of civic activity are unresponsive to prices than to think that aggregate levels of demand for consumer goods are unresponsive to prices.

¹³ See Paul A. Samuelson, *Foundations of Economic Analysis* 455 (enlarged ed. 1983).

preferences of people do not influence the aggregate level of civic acts, just as some Marxists mistakenly suppose that the preferences of people do not influence the aggregate level of commodities.

To illustrate the problem of pure skepticism, consider that Oliver Wendell Holmes urged legal scholars to reason from the viewpoint of a "bad man" without respect for law.¹⁴ Insofar as the bad man obeys the law, he does so for instrumental reasons. Thus the bad man treats law as "external" in the sense of being outside of his own values. Economic models of law typically accept the "bad man" approach and add an element to it: rationality. A bad man who is rational decides whether or not to obey the law by calculating his own benefits and costs, including the risk of punishment. The rational bad man breaks the law whenever the gain to him exceeds the risk of punishment. Law and economics scholars typically make the rational bad man into the decision-maker in their models. For the bad man, law is a constraint and not a guide.

In reality, society includes bad people and good citizens, as well as many citizens who are in between. Figure 3 shows how the effects of a law depend on the distribution of internalized values among citizens. Officials should not proceed by making laws only for bad people, because the response of good people also determines the laws' effects.

Instead of making the indefensible claim that internalization makes no difference to civic acts, skeptics do better by asserting that law makes no difference to the values internalized by citizens. To be precise, if law shifts the net cost curve more easily than the willingness-to-pay curve, then officials can influence citizens more effectively by manipulating costs rather than manipulating internalized values. In these circumstances, officials should proceed as if "tastes are given." Evaluating this form of skepticism requires an account of the causes of internalized values, which is my next topic.

IV. CAUSES OF INTERNALIZATION

Why do some people have tastes that others lack? Without an answer to this question, explaining different behavior by different

¹⁴ Oliver Wendell Holmes, *The Path of the Law*, 10 Harv. L. Rev. 457, 459 (1897), reprinted in 78 B.U. L. Rev. 699, 701 (1998).

tastes is a tautology, like asserting that opium makes people drowsy because of its "dormative power." Psychologists have studied non-rational and irrational influences on values, and philosophers have long disputed about rational values. In the past, however, economists made little effort to explain tastes, and some economists, such as Gary Becker and George Stigler, praised this reluctance as a methodological virtue linked to scientific rigor.¹⁵ A small but growing economics literature uses economic models of rationality to explore endogenous preferences,¹⁶ but economics still lacks an accepted method to investigate the internalization of values. Economic theory is only beginning a sustained inquiry into the question, "Where do tastes come from?"

I will extend the core concepts of economics to predict when rational people will internalize social norms. The theory of exchange states abstractly the conditions under which trade is mutually beneficial. Similarly, I begin with the conditions under which the internalization of social norms is mutually beneficial.

Business, politics, love, and war cause people to form relationships with each other. These relationships create opportunities for mutual benefit from cooperation and also opportunities for people

¹⁵ See George J. Stigler & Gary S. Becker, *De Gustibus Non Est Disputandum*, 67 *Am. Econ. Rev.* 76, 76-77 (1977).

¹⁶ See, e.g., Gary S. Becker, *Accounting for Tastes* 4 (1996); David Gauthier, *Morals By Agreement* (1986); Samuel Bowles, *Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions*, 36 *J. Econ. Literature* 75 (1998); Avinash Dixit & Victor Norman, *Advertising and Welfare*, 9 *Bell J. Econ.* 1 (1978); Avinash Dixit & Victor Norman, *Advertising and Welfare: Another Reply*, 11 *Bell J. Econ.* 753 (1980); Avinash Dixit & Victor Norman, *Advertising and Welfare: Reply*, 10 *Bell J. Econ.* 728 (1979); Jon Elster, *More Than Enough*, 64 *U. Chi. L. Rev.* 749 (1997) (reviewing Gary S. Becker, *Accounting for Tastes* (1996)); Robert Frank, *If Homo Economicus Could Choose His Own Utility Function, Would He Want One with a Conscience*, 77 *Am. Econ. Rev.* 593, 593-94 (1987); Peter J. Hammond, *Changing Tastes and Coherent Dynamic Choice*, 43 *Rev. Econ. Stud.* 159 (1976); Michael Hechter, *The Role of Values in Rational Choice Theory*, 6 *Rationality & Soc'y* 318 (1994); Maurice McManus, *Social Welfare Optimization with Tastes as Variables*, 101 *Weltwirtschaftliches Archiv* 114 (1978); Bezalel Peleg & Menahem E. Yaari, *On the Existence of a Consistent Course of Action When Tastes Are Changing*, 40 *Rev. Econ. Stud.* 391 (1973); Robert A. Pollak, *Habit Formation and Long-Run Utility Functions*, 13 *J. Econ. Theory* 272 (1976); Stigler & Becker, *supra* note 15; Robert Sugden, *Contractarianism and Norms*, 100 *Ethics* 768 (1990); Carl Christian von Weizsäcker, *Notes on Endogenous Change of Tastes*, 3 *J. Econ. Theory* 345 (1971); Menahem E. Yaari, *Endogenous Changes in Tastes: A Philosophical Discussion*, in *Decision Theory and Social Ethics* 59 (Hans W. Gottinger & Werner Leinfellner eds., 1978); Bruce L. Benson, *Endogenous Morality* (1997) (unpublished manuscript, on file with Virginia Law Review Association).

to exploit each other. People, consequently, care about the extent to which their partners internalize norms. According to Figure 3, a person who internalizes a norm is willing to pay to conform to it. In many activities, people prefer partners who are willing to pay a lot to conform to moral norms. Internalizing a norm, consequently, increases a person's opportunities.

To illustrate by analogy, signing an enforceable contract raises the cost of breach, thereby committing the promisor to perform. The promisor's commitment enables the promisee to rely on the promise, which makes the cooperation more profitable. So promisor and promisee both benefit from the former's legal commitment.¹⁷ Similarly, internalizing a norm is a moral commitment. Moral commitment increases the price that the person would pay to do what is right. People have more trust in the morally committed person. Trust facilitates cooperation and makes the relationship more beneficial to both parties.

This analysis assumes that people can make moral commitments. How do they morally commit? Is commitment psychological or rational? Does commitment come from habits or beliefs? Is a purely rational person morally uncommitted? The plethora of psychiatrists and self-help books testifies to the fact that we cannot automatically change ourselves by wanting to be different. Changing ourselves is a difficult technical problem, and I will not offer a theory to solve it in this Paper. Instead, I assume the existence of a technology for preference change without explaining it. In other words, I assume that people can change their preferences at some cost.

I have explained why people often prefer partners who have internalized morality. To act on this preference, people must observe the preferences of others. "Revealed preference theory" in economics examines rigorously the assumptions under which a person's choices reveal his preferences.¹⁸ Inferring moral commit-

¹⁷ For elaboration, see the agency theory of contracts, developed in Robert Cooter & Thomas Ulen, *Law and Economics* 184-88 (2000).

¹⁸ See Paul Samuelson, *Numerical Representation of Ordered Classifications and the Concept of Utility*, *RES* 5 (1938). For critical appraisals, see Daniel M. Hausman & Michael S. McPherson, *Economics, Rationality, and Ethics*, in *The Philosophy of Economics* 252, 256-59 (Daniel M. Hausman ed., 2d ed. 1994); Thomas Russell & Richard Thaler, *The Relevance of Quasi Rationality in Competitive Markets*, 75 *Am. Econ. Rev.* 1071, 1071-75 (1985); Amartya Sen, *Maximization and the Act of Choice*,

ments from behavior is difficult, especially among strangers. The moral commitments of strangers are opaque, but, as people get to know each other better, their moral commitments become translucent. In the case of very intimate relationships, such as mother-daughter, moral commitments can become transparent. In general, as relationships grow longer and deeper, moral commitments move from left to right on the following scale of observability.

Figure 6: Scale of Observability

Given that preferences are translucent and people care about the preferences of their partners, a person's preferences affect his opportunities. For example, if employers have some insight into the traits of employees, and if most employers prefer honest employees, then an honest person will enjoy more employment opportunities on average than a dishonest person. The dependence of opportunities on preferences gives a person an incentive to change his preferences. If a dishonest youth wants more opportunities for employment, for example, he might become honest.

I want to characterize the conditions that destabilize preferences and cause them to change. Identifying these conditions will help to predict when people will internalize values. According to the preceding discussion, people tend to change their preferences when doing so causes a sufficiently large increase in their opportunities. How large is "sufficiently large"? I will use Figure 7 to answer this question. The axes represent consumption of the first and second goods, and the lines represent indifference curves. Assume that the actor has preference U_1 , which results in opportunities such that (x_1, y_1) is the feasible combination of the goods yielding the most satisfaction. Assume that the actor can change his preference from U_1 to U_2 . The set of points labeled "Pareto improvement" contains

65 *Econometrica* 745, 745-46 (1997); Amartya K. Sen, *Rational Fools: A Critique of the Behavioural Foundations of Economic Theory*, 6 *Phil. & Pub. Aff.* 317, 323-26 (1977).

all combinations of the first and second good that yield higher satisfaction than (x_i, y_i) relative to preferences U_1 and U_2 .

Figure 7: Pareto Self-Improvement

Welfare economics calls a change that is better with respect to the preferences of everyone affected a “Pareto improvement.” By analogy, I call a change in preferences that leaves the decision-maker better off with respect to his original preferences and his final preferences a “Pareto self-improvement.”¹⁹

People will tend to change their preferences when doing so increases their opportunities sufficiently to increase their satisfaction relative to their initial preferences and their final preferences. I will explain this prediction. The preferences of people have structure. To illustrate, a person’s choices might reveal a preference for milk over soda. Behind this choice might lay a preference for health over infirmity. The first-order preference is for milk and the second-order preference is for health. Behind our particular preferences lie

¹⁹ I analyze this concept graphically in Cooter, *Expressive Law*, *supra* note 10, at 599; Cooter, “Bad Man” of Holmes, *supra* note 11, at 923.

more general, abstract preferences. The existence of different orders of preferences provides a potential explanation for how a rational person would choose his own preferences. In general, higher-order preferences can provide a guide for changing lower-order preferences. If a person's lower-order preferences determine his opportunities, then he might choose his lower-order preferences so that the resulting opportunities maximize his higher-order preferences.

Theorists have tried to identify a supreme value that orders all specific choices. Traditional candidates among philosophers include pleasure and happiness. Thus a hedonist would choose honesty over dishonesty, provided that honesty increased his pleasure. Similarly, a eudaemonist would choose honesty over dishonesty provided that honesty increased his happiness. Economists sometimes focus on wealth as supreme, so "homo economicus" might choose honesty over dishonesty, provided that honesty increases his wealth. Political theorists sometimes focus on power as supreme, so "political man" might choose honesty over dishonesty provided that honesty increases his power.

Compared to philosophers, ordinary people give relatively little thought to their higher-order preferences. Many people are unsure about the extent to which they prefer, say, wealth to pleasure, or happiness to success. These people need some guidance as they discover their higher order preferences by trial and error. Pareto self-improvements provide a guide to people who are uncertain about their higher order preferences. For example, if learning diligence in school promotes happiness and wealth, then a student who remains unsure about the relative importance of happiness and wealth still has a reason to learn diligence. After making the change and learning diligence, the opportunities enjoyed by the student are better than before relative to both his new and old preferences. Whether the actor evaluates the outcome using his new or old preferences, he will not regret making a Pareto self-improvement. This fact provides a reason to make the change for someone with imperfect information concerning his higher order preferences.

V. INTERNALIZATION AND LAW

I have explained that self-interested people change their preferences to increase their opportunities. Now I will relate this causal mechanism for changing preferences to the law's ability to make citizens internalize values. To induce people to internalize civic values, the state must reward citizens for having civic virtue. For this purpose, officials have honors, awards, and praise, as well as dishonor, punishments, and blame. To reward people for having civic virtue, however, the state must infer character from behavior. Inferring character from behavior requires intimate knowledge of the person. The officials in large states are remote from most citizens. The character of each citizen is relatively opaque to state officials. Consequently, officials lack the information needed to reward people for acquiring civic virtue. Instead of rewarding or punishing character, state law mostly rewards or punishes acts.

Compared to the state, people in intimate relationships with each other are relatively good at inferring character from behavior. Preferences are relatively transparent among intimates. Consequently, the primary influences on character are intimate relationships such as families, friends, and colleagues. Given these facts, the state will have little success instilling civic virtue. Instead, the state should rely on family, friends, and colleagues to instill civic virtue in each other. Insofar as family, friends, and colleagues prefer relationships with civic-minded people, individuals have an incentive to cultivate civic virtue. Civic acts can become signals for the possession of the moral traits that people seek in partners for cooperative ventures.

The primary way to prompt people to instill civic virtue in each other is by aligning law with morality. When law aligns with morality, individuals who cultivate morality necessarily acquire civic virtue. Consequently, the law enlists the force of internalized morality to achieve the ends of the state. To illustrate this mechanism, I will reinterpret Figure 5. Assume that revising a law to align it more closely with morality causes the upward shift in the willingness-to-pay curve. In other words, citizens are more willing to do their civic duty after the state grounds the duty in morality. This change shifts up the willingness-to-pay curve as depicted in Figure 5.

I have explained that people change their preferences to increase their opportunities, especially in intimate relationships. I

have argued that state officials, being remote from citizens, have limited ability to change citizens' preferences. Perhaps state officials can develop ways to change people in special circumstances, such as rehabilitating wrongdoers by "reintegrative shaming."²⁰ Or perhaps state officials can learn to use the techniques of commercial advertising. Commercial advertising often associates products with traits desired by partners in relationships. Commercial advertising often suggests that, say, men who consume a certain product are more attractive to women. Men respond to such advertising by changing their preferences in favor of the advertised product. Compared to businesses, modern states advertise relatively little. California spends large sums to register voters, but the state does not develop television advertisements on the theme, "Women prefer men who vote." Similarly, California spends large sums to catch speeding motorists, but the state does not develop television advertising on the theme, "Women prefer men who drive carefully." These themes probably make more sense than successful commercial advertising on such themes as "Women prefer men who eat potato chips."

I have discussed the possibility that people respect laws that align with morality. Now I consider the possibility that some people's morality requires them to respect law. In other words, I assume that some people believe that they ought to obey the law because it is law. These people align their morality with law. In *The Concept of Law*, H.L.A. Hart points out that many judges view law as a guide whose instructions they should follow.²¹ Rather than treating law as an external constraint, these judges value following the law. In developing this line of thought, Ronald Dworkin's influential early work asserted that judges could arrive at the right answer in a hard case by constructing the best synthesis of current legal rules and values.²² In contrast, other scholars favor the view

²⁰ As developed by John Braithwaite, the wrongdoer must confront his victims before his relatives, thus producing a combination of shame and forgiveness that reintegrates the wrongdoer into community life on the condition that he try to change. See John Braithwaite, *Crime, Shame and Reintegration* 55 (1989).

²¹ See Hart, *supra* note 6, at 141-42.

²² Dworkin argues that the great judge, Hercules, will find one right answer to every legal question, and the best synthesis of existing law gives the right answer. In this view, the law includes values as well as rules. Hercules, however, does not draw on his own personal or political values to decide cases. Dworkin, *supra* note 6, at 125-26. Later Dworkin modified his theory to allow some role for the personal and political

that independent judges express their own political vision in deciding hard cases.²³ Insofar as judges decide cases in light of their own political views, respect for law does not determine their decisions.

What actually motivates judges? If a judge decides cases strictly based on legal values, then the judge's political philosophy would not predict his decisions. In fact, social scientists can often predict judicial decisions from the political philosophy of judges.²⁴ Furthermore, U.S. federal judges behave differently according to various strategic considerations, such as the composition of the

values of the judge. See Ronald Dworkin, *Law's Empire* 255–56 (1986), where Dworkin argues that the ideal of law demands principled justification for the exercise of force by the state, and the rule of law requires judges to reach decisions that not only are consistent in principle with past legal authorities, but also characterize the relevant authorities in the best moral light.

²³ See Richard A. Posner, *What Do Judges and Justices Maximize? (The Same Thing Everybody Else Does)*, 3 *Sup. Ct. Econ. Rev.* 1, 39–40 (1993); see also Cooter & Ulen, *supra* note 17, at 408 (suggesting that the independence of American judges allows them to decide cases based on their own visions of right and wrong). Schauer emphasizes that the reputation of modern U.S. judges rests increasingly on the political outcomes of their decisions, so the prudential incentives to law-following are all we need to explain about the concept of law. See Frederick Schauer, *Autonomy, Judicial Activism and Judicial Review in Australian Democracy*, in *Legal Positivism and Legal Autonomy* (1998); Frederick Schauer, *Against Normativity*, Conference on Norms and Incentives, Australian National University Research School of Social Sciences (July 20–21, 1998).

²⁴ See Jeffrey A. Segal & Harold J. Spaeth, *The Supreme Court and the Attitudinal Model* 221–31 (1993); Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 *Va. L. Rev.* 1717, 1717–19 (1997) [hereinafter Revesz, *Environmental Regulation*]; Richard L. Revesz, *Ideology, Collegiality, and the D.C. Circuit: A Reply to Chief Judge Harry T. Edwards*, 85 *Va. L. Rev.* 805, 805 (1999) [hereinafter Revesz, *Ideology*]; Mario Bergara et al., *Judicial Politics and the Econometrics of Preferences, The Economics of Courts*, Conference sponsored by John M. Olin Center for Law, Economics, and Business (1999); F. Andrew Hanssen, *Independent Courts and Administrative Agencies: An Empirical Analysis of the States, The Economics of Courts*, Conference sponsored by John M. Olin Center for Law, Economics, and Business (1999); Richard L. Revesz, *Judges as Strategic Ideologues?: A Case Study of Challenges to Health-and-Safety Regulations in the D.C. Circuit, "The Economics of Courts,"* Conference sponsored by John M. Olin Center for Law, Economics, and Business (1999) [hereinafter Revesz, *Judges as Strategic Ideologues?*]; Edward P. Schwartz, *The Proliferation of Concurring Opinions on the U.S. Supreme Court: Politics Killed the Norm*, American Law and Economics Association (1997); Joseph L. Smith & Emerson H. Tiller, *The Strategy of Judging: Evidence from Administrative Law, The Economics of Courts*, Conference sponsored by John M. Olin Center for Law, Economics, and Business (1999); Joseph L. Smith & Emerson H. Tiller, *The Strategy of Judging: An Empirical Assessment* (American Law and Economics Association 1997) [hereinafter Smith, *The Strategy of Judging*]. But see Harry T. Edwards, *Collegiality and Decision Making on the D.C. Circuit*, 84 *Va. L. Rev.* 1335, 1358–62 (1998) (arguing that collegiality moderates the effect individual judges' views have on decisions).

panel on which they sit.²⁵ Empirical evidence thus supports the theory that judges respond to values outside the narrow confines of the law. This evidence refutes the claim that independent judges act purely from respect for law. However, the unexplained residual remains large in these statistical studies. Future research will have to determine whether or not respect for the law explains the currently unexplained residual.

To illustrate how respect for law among citizens might influence their behavior, I will reinterpret Figure 5 once again. Assume that officials identify a moral obligation and raise it to the level of a legal obligation. For citizens who place no intrinsic value on obeying the law, the enactment of the law does not change their willingness to do their duty. For citizens who intrinsically value obeying the law, however, the enactment of the law increases their willingness to do their duty. Consequently, the willingness-to-pay curve shifts up in Figure 5.

These two mechanisms for support of law by citizens—alignment of law with morality and respect for law—are allegedly interrelated. According to some philosophers, when citizens reflect upon the state, they often evaluate its performance against standards of justice. If the state performs well relative to these standards, then reflective citizens conclude that the law deserves respect. Once citizens respect the law, they obey it habitually in their daily lives. Thus a just state achieves stability by generating its own support among reflective citizens.²⁶

²⁵ See Revesz, *Environmental Regulation*, supra note 24, at 1764; Revesz, *Ideology*, supra note 24, at 834–36; Revesz, *Judges as Strategic Ideologues?*, supra note 24; Schwartz, supra note 24; Smith, *The Strategy of Judging*, supra note 24; cf. Edwards, supra note 24, at 1358–62 (discussing the pattern of consensus on the D.C. Circuit in light of collegiality); Eric Rasmusen, *Judicial Legitimacy as a Repeated Game*, 10 *J.L. Econ. & Org.* 63, 81 (1994) (discussing the strategy of current judges influencing and constraining future judges); Eli Salzberger & Paul Fenn, *Judicial Independence: Some Evidence from the English Court of Appeal*, 42 *J.L. & Econ.* 831 (1999) (suggesting that how often a British judge agrees with the party line has an influence on her chances for promotion to a higher court).

²⁶ Rawls uses this line of thought to argue for the stability of states organized according to his principles of justice. See John Rawls, *A Theory of Justice* 434–41 (rev. ed. 1999).

CONCLUSION

With civic acts and commodities, tastes and costs simultaneously determine the equilibrium quantity. The skeptical claim that internalized values do not matter reduces to the false claim that costs alone determine the equilibrium. Instead of making the indefensible claim that internalization makes no difference to civic acts, skeptics do better by asserting that law makes no difference to the values internalized by citizens. To be precise, if law shifts the net cost curve more easily than the willingness-to-pay curve, then officials can influence citizens more effectively by manipulating costs rather than by manipulating internalized values.

Evaluating this weaker form of skepticism requires an account of the causes of internalized values. People especially change their preferences to increase their opportunities. Specifically, people internalize morality to improve their opportunities for cooperating with others. Moral commitments are translucent in intimate relationships and opaque in remote relationships. Since officials have remote relationships with citizens in modern states, the state has little power to induce people to make moral commitments. The law must, consequently, reward and punish acts, not the actor's character. Instead of promoting civic virtue directly, the state must align law with social norms. Enlisting preexisting morality in the service of the state helps solve agency problems that plague government.

* * *