MODELS OF MORALITY IN LAW AND ECONOMICS: SELF-CONTROL AND SELF-IMPROVEMENT FOR THE "BAD MAN" OF HOLMES

Robert Cooter*

INTRODUCTION

Holmes commended scholars to analyze law from the viewpoint of a "bad man."[1] Who is he? In Milton's Paradise Lost, Satan says

* Herman F. Selvin Professor of Law, Boalt Hall School of Law, University of California at Berkeley.

[1] See Oliver Wendell Holmes, The Path of the Law, 10 HARV. L. REV. 457, 459 (1897). The full passage reads as follows:

> If you want to know the law and nothing else, you must look at it as a bad man, who cares only for the material consequences which such knowledge enables him to predict, not as a good one, who finds his reasons for conduct, whether inside the law or outside of it, in the vaguer sanctions of conscience.

Id.

"Evil be thou my Good."[2] The worst villains do bad for its own sake. The poet Coleridge called such behavior "motiveless malignity," because evil serves no further purpose beyond itself.[3] Such lofty evil, however, seems rare relative to everyday wrongdoing. In The Fall of Rome, Auden writes:

> Fantastic grow the evening gowns;
> Agents of the Fisc pursue
> Absconding tax-defaulters through
> The sewers of provincial towns.[4]

Auden thus attributes an empire's collapse to petty selfishness and materialism.

If Holmes had in mind Auden's everyday wrongdoing, and not Milton's lofty evil, then he prescribed the research strategy of law and economics. The virtuous prefer good, villains prefer bad, and rational actors in economics prefer themselves. Instead of obeying or disobeying the law for its own sake, the rational actor in economics treats law as an obstacle or an instrument,

---

[2] JOHN MILTON, PARADISE LOST bk. IV, l. 110, at 86 (Merritt Y. Hughes ed., Odyssey Press 1962) (2d rev. ed. 1674).

[3] SAMUEL TAYLOR COLERIDGE, Notes on Othello, in THE LITERARY REMAINS OF SAMUEL TAYLOR COLERIDGE 255, 260 (Henry Nelson Coleridge ed., AMS Press 1967) (1836) (describing Iago's soliloquy as "the motive-hunting of the motiveless malignity").

[4] W.H. Auden, The Fall of Rome, in 2 THE OXFORD ANTHOLOGY OF ENGLISH LITERATURE 2110, 2110 (Frank Kermode & John Hollander eds., 1973).

not a value. The success of the economic analysis of law proves the fruitfulness of this research strategy.[5]

Instead of praise, however, this paper offers critique and extension. The social complexity that law regulates increases with population, technical knowledge, and production. To cope with complexity, economists typically prescribe decentralization. Decentralized law works best when spontaneous obedience and private enforcement supplement state coercion. Internalized morality prompts spontaneous obedience and perfects private enforcement.[6] Thus minimizing state coercion maximizes reliance upon internalized morality.

Developing a theory of morality to anchor decentralized law requires extending models beyond the bad man. This paper develops two fundamental ideas of morality: self-control and self-improvement. I will explain how law harnesses and strengthens self-control, and also how the law changes people by creating opportunities for self-improvement.

Wrongdoing often yields an immediate benefit and risks

---

[5] The success of conventional law and economics is proved and charted in William M. Landes & Richard A. Posner, The Influence of Economics on Law: A Quantitative Study, 36 J.L. & ECON. 385 (1993).

[6] For models demonstrating how internalized morality by some people affects aggregate behavior through spontaneous obedience and private enforcement, see Robert Cooter, Normative Failure Theory of Law, 82 CORNELL L. REV. 947, 953-59 (1997) [hereinafter Cooter, Normative Failure Theory].

[7] Mill developed this theory in Utilitarianism:

future punishment. In balancing these considerations, a rational person must discount punishment for futurity and uncertainty. People who lack self-control apply different discount rates depending upon their mood. When their mood causes them to discount too deeply, wrongdoing seems more attractive. In contrast, people with self-control apply the same discount rate regardless of their mood. Thus self-control reduces the variance in subjective discount rates, which reduces the probability of spontaneous wrongdoing.

This fact has implications for law and policy. For example, differences in variance of discount rates between young and old people imply differences in optimal punishments. For imprudent youth, relatively mild punishments applied with high probability deter optimally. For prudent adults, relatively severe punishments applied with low probability deter optimally.

Acquiring self-control changes preferences. I introduce a novel concept to explain how and why rational people change their preferences. A change in preferences can cause a change in opportunities. For example, an employer may promote an employee who acquires more self-control and becomes more reliable. Define a "Pareto self-improvement" as a change in preferences that creates a feasible alternative preferred by the old preferences and the new preferences. For example, becoming more reliable can increase earnings enough to make the person better off as measured by the old preferences (unreliable) and the new

preferences (reliable). I contend that a rational person will make Pareto self-improvements whenever the opportunity arises.

Punishment creates opportunities for Pareto self-improvements. To illustrate, children dissemble poorly, so the best way for a child to appear to be honest is actually to be honest. Parents often punish dishonest children by denying them freedom and responsibility. In such a social environment, the advantages of honesty to the child may outweigh the disadvantages. Thus the threat of punishment can tip the balance of net benefits in favor of improving oneself.

State sanctions, which some theorists consider law's essence,[7] can produce similar incentives for adults to improve their character as parental punishment produce for children. I will show that legal punishments create opportunities for Pareto self-improvements where none would otherwise exist.

I. A BRIEF REVIEW OF THE LITERATURE

I will briefly review the connection between legal scholarship on morality and economic analysis. Proponents of decentralized law have long admired social norms because they arise spontaneously outside the state.[8] Legal scholars, however,

---

[8] See, e.g., DAVID FRIEDMAN, THE MACHINERY OF FREEDOM 105-08 (1973) (arguing that a fundamental decentralization of large American cities would lead to more responsive public services); F.A. HAYEK, THE ROAD TO SERFDOM 83-87 (5th ed., University of Chicago Press 1994) (1944) (explaining that a decentralized system of government aids in the development of individual personalities); BRUNO LEONI, FREEDOM AND THE LAW 135-41 (1961) (distinguishing the "common will" from majoritarian decision-making on the theory

underestimated the effects of social norms until empirical research proved that they control behavior in spite of law. To illustrate, American businesses remain rationally ignorant of the legal consequences of the contracts that they sign,[9] social norms affect drunk driving enforcement in spite of legislation,[10] Peruvian businesses systematically break the law to circumvent

---

that the common will "emerges from the collaboration of all the people concerned, without any recourse to group decisions and decision groups" (emphasis omitted)); Paul H. Rubin, Growing a Legal System in the Post-Communist Economies, 27 CORNELL INT'L L.J. 1, 7-9 (1994) (arguing that post-cold war Eastern European countries should rely on "common-law like processes" to develop their respective laws of contracts).

[9] See Stewart Macaulay, Non-Contractual Relations in Business: A Preliminary Study, 28 AM. SOC. REV. 55, 60 (1963) (concluding that "many, if not most, exchanges reflect no planning, or only a minimal amount of it, especially concerning legal sanctions and the effect of defective performance.")

[10] See H. LAURENCE ROSS, DETERRING THE DRINKING DRIVER 103 (1982) (concluding that "[c]hanges in behaviour resulting from changes in the certainty of the threat . . . are evanescent"); H. Laurence Ross, Deterrence-based Policies in Britain, Canada, and Australia, in SOCIAL CONTROL OF THE DRINKING DRIVER 64, 66, 70 (Michael D. Laurence et al. eds., 1988) (documenting the short-term effect of deterrence-based measures in Britain and Canada).

[11] HERNANDO DE SOTO, THE OTHER PATH 232 (June Abbott trans., Harper & Row 1988) (n.d.) (concluding that the Peruvian government is corrupt to the point that "legal institutions have ceased to provide the means to govern society and to live in it").

excessive regulations,[11] and small business financing in Taiwan often occurs outside of formal law.[12]

The demonstrated importance of social norms to law and the availability of analytical techniques from economics have caused a renaissance in legal scholarship on social norms. This is illustrated by studies on liability for harm caused by cattle,[13] land courts among tribal people,[14] merchant courts,[15] and many

---

[12] See Jane Kaufman Winn, Relational Practices and the Marginalization of Law: Informal Financial Practices of Small Businesses in Taiwan, 28 L. & Soc'y Rev. 193, 211-25 (1994) (observing that Taiwanese small business financing often occurs via unregulated mechanisms, including rotating credit associations, postdated checks, and "underground moneylenders").

[13] See Robert Ellickson, Order Without Law 52-64 (1991) (observing that rural Shasta County, California residents prefer to resolve their disputes "beyond the shadow of the law," and detailing "how the norms of neighborliness operate and how deviants who violate these norms are informally controlled"); id. at 156-83 (using game theory to analyze the "puzzle of cooperation").

[14] See Robert D. Cooter, Inventing Market Property: The Land Courts of Papua New Guinea, 25 L. & Soc'y Rev. 759, 761-65 (1991) (explaining that 97% of the land in Papua New Guinea is controlled by a vaguely-defined system of customary ownership); id. (using economic theory to conclude that courts, not legislatures, should take the lead in inventing new forms of market property that remain congenial to custom).

[15] See Lisa Bernstein, Merchant Law in a Merchant Court: Rethinking the Code's Search for Immanent Business Norms, 144 U. Pa. L. Rev. 1765, 1796-1820 (1996) [hereinafter Bernstein, Merchant Law] (comparing a private judicial system within the feed and grain industry with the Uniform Commercial Code's adjudicative philosophy, and concluding that the Code's use of immanent business norms leads to market inefficiency); Lisa Bernstein, Opting Out of the Legal System: Extralegal Contractual Relations in the Diamond Industry, 21 J. Legal Stud. 115, 145-53 (1992) [hereinafter Bernstein, Opting Out] (comparing the efficiency of extralegal enforcement mechanisms in the diamond industry with legal enforcement mechanisms).

subsequent studies.[16]

---

[16] For additional legal scholarship on norms, see Lisa Bernstein, Social Norms and Default Rules Analysis, 3 S. CAL. INTERDISC. L.J. 59 (1993); Dov Cohen & Joe Vandello, Meanings of Violence, 27 J. LEGAL STUD. (forthcoming June 1998); Robert D. Cooter, Decentralized Law for a Complex Economy: The Structural Approach to Adjudicating the New Law Merchant, 144 U. PA. L. REV. 1643 (1996) [hereinafter Cooter, Decentralized Law]; Cooter, Normative Failure Theory, supra note 6; Robert D. Cooter, The Theory of Market Modernization of Law, 16 INT'L REV. L. & ECON. 141 (1996); Robert C. Ellickson, Law and Economics Discovers Social Norms, 27 J. LEGAL STUD. (forthcoming June 1998); Richard L. Hasen, Voting Without Law?, 144 U. PA. L. REV. 2135 (1996); Dan M. Kahan, Social Meaning and the Economic Analysis of Crime, 27 J. LEGAL STUD. (forthcoming June 1998); Timur Kuran, Ethnic Norms and Their Transformation Through Reputational Cascades, 27 J. LEGAL STUD. (forthcoming June 1998); Lawrence Lessig, The New Chicago School, 27 J. LEGAL STUD. (forthcoming June 1998); Lawrence Lessig, The Regulation of Social Meaning, 62 U. CHI. L. REV. 943 (1995); Lawrence Lessig, Social Meaning and Social Norms, 144 U. PA. L. REV. 2181 (1996); Geoffrey P. Miller, The Song of Deborah: A Legal Economic Analysis, 144 U. PA. L. REV. 2293 (1996); Martha C. Nussbaum, "Whether from Reason or Prejudice": Taking Money for Bodily Services, 27 J. LEGAL STUD. (forthcoming June 1998); Randal C. Picker, Simple Games in a Complex World: A Generative Approach to the Adoption of Norms, 64 U. CHI. L. REV. 1225 (1997); Eric A. Posner, Law, Economics, and Inefficient Norms, 144 U. PA. L. REV. 1697 (1996) [hereinafter Posner, Law, Economics, and Inefficient Norms]; Eric A. Posner, The Legal Regulation of Religious Groups, 2 LEGAL THEORY 33 (1996); Eric A. Posner, Efficient Norms, in THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW (Peter Newman ed., forthcoming 1998); Eric A. Posner, Symbols, Signals, and Social Norms in Politics and the Law, 27 J. LEGAL STUD. (forthcoming June 1998); Richard A. Posner, Social Norms, Social Meaning, and Economic Analysis of Law: A Comment, 27 J. LEGAL STUD. (forthcoming June 1998); J. Mark Ramseyer, Learning to Love Japan: Social Norms and Market Incentives, 34 SAN DIEGO L. REV. 263 (1994); Alan Schwartz, Karl Llewellyn and the Origins of Contract Theory, in THE JURISPRUDENTIAL FOUNDATIONS OF CORPORATE AND COMMERCIAL LAW (Jody S. Kraus & Steven D. Walt eds., forthcoming 1999); Cass R. Sunstein, Social Norms and Social Roles, 96 COLUM. L. REV. 903 (1996); Elizabeth Anderson, Should Feminists Reject Rational Choice Theory? (1998) (unpublished paper, on file with Elizabeth Anderson, University of Michigan). In addition, two recent journal issues are devoted exclusively to this topic. See Symposium, Law, Economics, & Norms, 144 U. Pa. L. Rev. 1643

Formal analyses of social norms apply game theory.[17] One-

(1996); Symposium, Social Norms, Social Meaning, and the Economic Analysis of Law, 27 J. LEGAL STUD. (forthcoming June 1998).

[17] See, e.g., JACK HIRSHLEIFER, Evolutionary Models in Economics and Law: Cooperation versus Conflict Strategies, in ECONOMIC BEHAVIOUR IN ADVERSITY 211 (using game theory to trace the interplay of cooperative and conflicting motivations in the evolution of legal systems and economic norms); MICHAEL TAYLOR, THE POSSIBILITY OF COOPERATION (1987) (using game theory to examine the possibility of voluntary cooperation in the provision of public goods and in the solution of other collective action problems); EDNA ULLMAN-MARGALIT, THE EMERGENCE OF NORMS (1977) (using game theory to analyze coordination and inequality norms); Robert Sugden, Reciprocity: The Supply of Public Goods Through Voluntary Contributions, 94 ECON. J. 772 (1984) (using game theory to model the "voluntary sector"). For other game theory and economics papers on norms, see GARY S. BECKER, ACCOUNTING FOR TASTES (1996) [hereinafter BECKER, ACCOUNTING FOR TASTES]; GARY S. BECKER, A TREATISE ON THE FAMILY (enlarged ed. 1991) (1981); George A. Akerlof, Discriminatory, Status-based Wages among Tradition-oriented, Stochastically Trading Coconut Producers, 93 J. POL. ECON. 265 (1985) [hereinafter Akerlof, Discriminatory, Status-based Wages]; George A. Akerlof, A Theory of Social Custom, of Which Unemployment May Be One, 94 Q.J. ECON. 749 (1980) [hereinafter Akerlof, A Theory of Social Custom]; B. Douglas Bernheim, A Theory of Conformity, 102 J. POL. ECON. 841 (1994); B. Douglas Bernheim & Oded Stark, Altruism Within the Family Reconsidered: Do Nice Guys Finish Last?, 78 AM. ECON. REV. 1034 (1988); Timothy Besley, Nonmarket Institutions for Credit and Risk Sharing in Low-Income Countries, 9 J. ECON. PERSP. 115 (1995); Susil Bikchandani et al., A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades, 100 J. POL. ECON. 992 (1992); David de Meza & J.R. Gould, The Social Efficiency of Private Decisions to Enforce Property Rights, 100 J. POL. ECON. 561 (1992); Robert H. Frank, If Homo Economicus Could Choose His Own Utility Function, Would He Want One with a Conscience?, 77 AM. ECON. REV. 593 (1987) [hereinafter Frank, If Homo Economicus Could Choose]; Avner Greif et al., Coordination, Commitment, and Enforcement: The Case of the Merchant Guild, 102 J. POL. ECON. 745 (1994); Timur Kuran, Islamic Economics and the Islamic Subeconomy, 9 J. ECON. PERSP. 155 (1995); Paul R. Milgrom et al., The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs, 2 ECON. & POL. 1 (1990); David Romer, The Theory of Social Custom: A Modification and Some Extensions, 99 Q.J. ECON. 717 (1984); Michael Taylor, Cooperation, Norms, and Moral Motivation, 15 ANALYSE & KRITIK 70 (1993); Michael Taylor, Good Government: On

shot games with inefficient solutions may have efficient solutions when repeated.[18] This fact explains the tendency of small groups to develop efficient rules for cooperation among themselves, as demonstrated by cattle ranchers,[19] Chinese traders,[20] medieval merchants,[21] and modern merchant

---

Hierarchy, Social Capital, and the Limitations of Rational Choice Theory, 4 J. POL. PHIL. 1 (1996).

[18] See ROBERT AXELROD, THE EVOLUTION OF COOPERATION 169-91 (1984) (observing that "trial and error" learning processes eventually lead to efficient results); Drew Fudenberg & Eric Maskin, The Folk Theorem in Repeated Games with Discounting or with Complete Information, 54 ECONOMETRICA 533, 547-52 (1986) (demonstrating that individually rational outcomes can be sustained as the equilibrium payoffs of a finitely repeated game if the number of repetitions is sufficiently large).

[19] See ELLICKSON, supra note 13 at 167-83 (arguing that the ranchers of Shasta County, California maintain a set of norms that maximizes their aggregate welfare).

[20] See Janet T. Landa, The Political Economy of the Ethnically Homogenous Chinese Middleman Group in Southeast Asia: Ethnicity and Entrepeneurship in a Plural Society, in 1 THE CHINESE IN SOUTHEAST ASIA: ETHNICITY AND ECONOMIC ACTIVITIES 86, 95-101 (L.A. Peter Gosling & Linda Y.C. Lim eds., 1983) (arguing that Chinese middlemen improve contract certainty by personalizing exchange relations along kinship and ethnic lines); Janet Landa, A Theory of the Ethnically Homogenous Middleman Group: An Institutional Alternative to Contract Law, 10 J. LEGAL STUD. 349, 357-61 (1981) (observing that the informal social networks within a group of Chinese middlemen lead to efficient information screening and mobilizing devices).

[21] See Greif et al., supra note 17 at 762-71 (observing that strong merchant guild associations provided an efficient trading mechanism during the early middle ages); Milgrom et al., supra note 17 at 6-18 (explaining the importance of reputational information to the medieval Law Merchant system).

associations.[22] Research on property rights reveals variety and detail in the political arrangements by which small groups manage their assets.[23]

The utilitarianism of small groups applies to people who interact repeatedly with each other, such as the Berkeley Chess Club, but not to categories of people who seldom interact, such as chess players in California.[24] Furthermore, one group may develop norms that benefit its members by exploiting or

---

[22] See Bernstein, Merchant Law, supra note 15, at 1815-20 (concluding that the use of reputational bonds in a private judicial system within the feed and grain industry is a more efficient than the Uniform Commercial Code's use of immanent business norms); Bernstein, Opting Out, supra note 15, at 138-53 (observing that an extralegal enforcement system based on reputational bonds within the diamond industry is more efficient than the traditional legal system).

[23] See Thráinn Eggertsson, Analyzing Institutional Successes and Failures: A Millennium of Common Mountain Pastures in Iceland, 12 INT'L REV. L. & ECON. 423, 425-31 (1992) (arguing that the historical joint utilization of the mountain pastures of Iceland was economically rational); Robert C. Ellickson, Property in Land, 102 YALE L.J. 1315, 1341-44 (1993) (suggesting that small group ownership of common land is an efficient means of allocating risk); id. at 1388-94 (observing that small open-field villages historically tended to exploit efficiencies of scale and spread risks); Donald N. McCloskey, The Economics of Enclosure: A Market Analysis, in EUROPEAN PEASANTS AND THEIR MARKETS 123, 127-42, 151-60 (William N. Parker & Eric L. Jones eds., 1975) (explaining that open-field agriculture in England had become inefficient by the late 1700s); Donald N. McCloskey, The Persistence of English Common Fields, in EUROPEAN PEASANTS AND THEIR MARKETS, supra, at 73, 113-19 (arguing that the scattering phenomenon in English open fields was an efficient means of risk-spreading).

[24] See Eric A. Posner, The Regulation of Groups: The Influence of Legal and Nonlegal Sanctions on Collective Action, 63 U. CHI. L. REV. 133, 144-50 (1996) (explaining that groups tend to cooperate, whereas categories do not).

subordinating non-members.[25]

Economic analyses demonstrate a surprising level of efficiency in common law rules.[26] Attempts failed to explain this

---

[25] See Akerlof, Discriminatory, Status-based Wages, supra note 17 at 268-75 (modeling the economic effects of status-based discrimination in a hypothetical coconut-producing island society); Akerlof, A Theory of Social Custom, supra note 17 at 757-72 (explaining that a discriminatory social norm benefiting one group's members could override the group's purely economic interests); Richard H. McAdams, Cooperation and Conflict: The Economics of Group Status Production and Race Discrimination, 108 HARV. L. REV. 1003, 1063-74 (1995) (explaining the persistence of race discrimination and the stability of discriminatory norms); Posner, Law, Economics, and Inefficient Norms, supra note 16, at 1722-23 (observing that some negative externality-producing social norms persist even though the norms are socially inefficient).

[26] This fact is demonstrated in the succession of text books that focus on the economic analysis of the common law. See, e.g., **[RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 271-91 (5th ed. 1998) (asserting that economic principles transcend subject-matter distinctions, forming the roots of contract, property, and tort principles)]**; **[ROBERT COOTER & THOMAS ULEN, LAW AND ECONOMICS ____ (2d ed. 1997) (**parenthetical**)]**; A. MITCHELL POLINSKY, AN INTRODUCTION TO LAW AND ECONOMICS (2d ed. 1989) (examining law and economics from a normative perspective).

[27] The possibility that competitive adjudication of inefficient rules results in an economically efficient body of common law was first addressed in Paul H. Rubin, Why is the Common Law Efficient?, 6. J. LEGAL STUD. 51, 53-57 (1977). See also George L. Priest, The Common Law Process and the Selection of Efficient Rules, 6 J. LEGAL STUD. 65, 66-75 (1977) (expanding on Rubin's hypothesis to conclude that "[e]fficient rules 'survive' in an evolutionary sense because they are less likely to be relitigated and thus less likely to be changed, regardless of the method of decision"). For an argument that judges may as well decide in terms of efficiency, since they have no other criteria to use, see RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 99 (1st ed. 1972) (observing that "[i]n searching for a reasonably objective and impartial standard, . . . the judge can hardly fail to consider whether the loss was the product of wasteful uneconomical resource use").

puzzle by competitive adjudication or judicial motivation.[27] Social norms, however, might explain the paradox. If judges selectively enforce social norms, and social norms evolve towards efficiency, the common law could evolve towards efficiency as a consequence of social forces rather than legal forces.[28]

Other social scientists besides economists exert a diffuse influence on the new legal scholarship concerning social norms. Legal scholars have been influenced by both empirical and theoretical sociology.[29] Social psychologists have accumulated

---

[28] See Robert Cooter, Structural Adjudication and the New Law Merchant: A Model of Decentralized Law, 14 INT'L REV. L. & ECON. 215, 224-26 (1994) (observing that "local improvements lead to a global maximum on a concave surface," and concluding that "strategies that evolve into social norms in a free business community will be efficient in the absence of nonconvexities or spillovers to other communities" (emphasis omitted)).

[29] See, e.g., KRISTIN LUKER, TAKING CHANCES: ABORTION AND THE DECISION NOT TO CONTRACEPT (1975) (conducting an exhaustive empirical analysis of contraceptive risk taking in California following the liberalization of state abortion law); TOWARD A GENERAL THEORY OF SOCIAL CONTROL (Donald Black ed., 1984) (collecting essays that seek to establish a theoretical conception of social control as a dependent variable). For some examples of recent sociology scholarship, see ROBERT B. EDGERTON, SICK SOCIETIES (1992) (examining and challenging the "myth" of primitive harmony by pointing out that many societies failed to adopt socially efficient norms); Michael Hechter & Satoshi Kamazawa, Group Society and Social Order in Japan, 5 J. THEORETICAL POL. 455 (1993) (arguing that the complexity of Japanese social order derives from unintended mechanisms that lead to group solidarity); Toshio Yamagashi, The Provision of a Sanctioning System in the United States and Japan, 51 SOC. PSYCHOL. Q. 265 (1988) (reporting the results of an empirical cooperation/sanctioning experiment, and concluding that the Japanese subjects, who live in a society characterized by strong mutual monitoring and sanctioning, cooperate less in the absence of a sanctioning system than their more individualistic American counterparts).

impressive evidence that people obey the law more from internalized respect than from fear of punishment.[30]

The renaissance in legal scholarship on social norms, although vigorous, suffers from the inability of economics to comprehend normative commitment. Theories of endogenous preferences, which go back at least to Adam Smith,[31] have not

---

[30] See TOM R. TYLER, WHY PEOPLE OBEY THE LAW 57-68 (1990) (presenting empirical evidence to show that individual compliance flows from the individual's sense of right and wrong as well as a personal feeling of obligation to obey the law); see also id. at 67 (observing that the study "found little evidence of deterrence effects").

[31] See ADAM SMITH, THE THEORY OF MORAL SENTIMENTS pt. V, ch. II, at 200-11 (D.D. Raphael & A.L. Macfie eds., Clarendon Press 1976) (6th rev. ed. 1690) (remarking on "the [i]nfluence of [c]ustom and [f]ashion upon [m]oral [s]entiments").

[32] See BECKER, ACCOUNTING FOR TASTES, supra note 17 at 18-19 (observing that "[t]he endogeneity of preferences makes it difficult to determine whether preferences or opportunities are responsible for particular economic outcomes"); ALBERT O. HIRSCHMAN, Against Parsimony: Three Easy Ways of Complicating Some Categories of Economic Discourse, in RIVAL VIEWS OF MARKET SOCIETY 142, 157-59 (1986) (concluding that traditional economic analyses fail to account for societal preferences); Michael Hechter, The Role of Values in Rational Choice Theory, 6 RATIONALITY & SOC'Y 318, 318-21 (1994) (criticizing traditional economic theory for failing to account for societal values); Richard H. Thaler & H.M. Shefrin, An Economic Theory of Self-control, 89 J. POL. ECON. 392, 393 (1981) (analyzing the concept of self-control using an agency-based theory, as opposed to "reliance on ad hoc explanations in which transaction costs, taxes and income effects play a major role"); Carl Christian von Weizsäcker, Notes on Endogenous Change of Tastes, 3 J. POL. ECON. THEORY 345, 345-46 (1971) (observing that "the overwhelming majority [of economists] took the attitude that it is not their business to be concerned with . . . changes of taste" and presenting a model of consumer behavior as a matter of personal taste). For two early attempts to incorporate endogenous preferences into conventional economic theory, see Robert A. Pollak, Habit Formation and Long-Run Utility Functions, 13 J. ECON. THEORY 272 (1976); Menahem E.

flourished in economics.[32] Microeconomics marginalizes morality by treating it as an exogenous taste or a side constraint upon optimizing behavior.[33] Thus the theory of cooperative games,

---

Yaari, Endogenous Changes in Tastes: A Philosophical Discussion, in DECISION THEORY AND SOCIAL ETHICS 59 (Hans W. Gottinger & Werner Leinfellner eds., 1976).

[33] The significance of the difference between morality as a preference and a constraint is explored in Matthew Rabin, Moral Preferences, Moral Constraints, and Self-serving Biases (Aug. 1995) (unpublished working paper No. 95-241, on file with the Berkeley Department of Economics) [hereinafter Rabin, Moral Preferences].

[34] To illustrate, the classic textbook on game theory devotes a chapter to cooperative games, see R. DUNCAN LUCE & HOWARD RAIFFA, GAMES AND DECISIONS 114-54 (1957), whereas one of the best modern books omits it, see DREW FUDENBERG & JEAN TIROLE, GAME THEORY (1991).

which requires normative commitments from players, languishes while the theory of non-cooperative games flourishes.[34]

Excluding cooperation from game theory favors purity over reality. Experimental evidence indicates the pervasiveness of cooperation in spite of the requirements of narrow self-interest.[35] Players who "irrationally" cooperate often gain an advantage in competition with narrowly instrumental players, thus straining the definition of rationality.[36] In experimental economics, the initial discovery of the resilience of moral

---

[35] See, e.g., Elizabeth Hoffman & Matthew L. Spitzer, Entitlements, Rights, and Fairness: An Experimental Examination of Subjects' Concepts of Distributive Justice, 14 J. LEGAL STUDIES 259, 281-84 (1985) (interpreting cooperation game data to conclude that the test subjects generally conformed to a Lockean "just deserts" theory, rather than a utilitarian "self-interest" theory); Elizabeth Hoffman et al., Preferences, Property Rights, and Anonymity in Bargaining Games, 7 GAMES & ECON. BEHAV. 346, 370-71 (1994) (concluding that fairness concerns motivated ultimatum and dictator game test subjects to disregard their individual self-interests, in spite of experimental anonymity); **[Kevin McCabe et al., Intentionality Detection and "Mindreading": Why Does Game Form Matter? __ (Apr. 1998) (unpublished working paper, on file with the Economic Science Laboratory, University of Arizona) (**parenthetical**)].**

[ER -- professor Vernon Smith of the Univ. of Ariz. offered to send a copy of the McCabe paper.]

[36] See AXELROD, supra note 18 at 175-76 (reporting that a cooperation-based computer model won the first two rounds of a prisoner's dilemma tournament over the competing entries of several professional game theorists); ROBERT H. FRANK, PASSIONS WITHIN REASON 67-70 (1988) (eschewing traditional economic notions of rationality in favor of a commitment-based theory in which individuals "value[] trustworthiness for its own sake"); Frank, If Homo Economicus Could Choose, supra note 17 at 601-03 (suggesting that rational "people will often refrain from cheating not because they fear being caught, but because cheating simply makes them feel bad").

commitment has yielded to progressive refinements that explain individual commitments.[37] Promising new developments in analyzing morality come from evolutionary models, in which normative commitment flourishes to the extent that a competitive environment rewards it.[38]

Unlike economics, philosophy often treats morality as rational, but rationality in non-utilitarian philosophy hardly resembles economic rationality.[39] Some moral philosophy seems to

---

[37] See, e.g., Hoffman & Spitzer, supra note 35 at 286-89 (maintaining that the presence of laboratory observers did not affect participant choices in a cooperation game); Hoffman et al., supra note 35 at 370-71 (retreating from the prior Hoffman & Spitzer premise by demonstrating that subject anonymity with respect to the observer resulted in lower initial offers in a dictator game).

[38] See, e.g., Abhijit Banerjee & Jörgen W. Weibull, Evolution and Rationality: Some Recent Game-Theoretic Results, in 2 ECONOMICS IN A CHANGING WORLD 90, 107-09 (Beth Allen ed., 1996) (summarizing current game theory approaches to evolution and rationality, and concluding that the combined approaches "provide[] some support for the rationalistic approach of non-cooperative game theory").

[39] Systematic western philosophy is often traced to Plato, whose Republic inquires into the rational basis of justice. See generally PLATO, THE REPUBLIC bks. I-II, *327a-*367e, at 3-45 (Allan Bloom trans., 1968) (n.d.) (developing the basic concept of "justice" and considering the relationship between justice and wealth maximization). The recent magisterial book by Rawls continues that inquiry. See JOHN RAWLS, A THEORY OF JUSTICE 142-50 (1971) (stressing that rational actors "try to acknowledge principles which advance their system of ends as far as possible" without regard to the comparative gains of others). Theories of rational morality that reject utilitarian reasoning often draw upon Kant. See, e.g., THOMAS NAGEL, THE POSSIBILITY OF ALTRUISM 13-17 (1970) (drawing on Kant's conceptions of ethical motivation and self-conception to form the basis of a theory of altruistic motivation).

[40] See DAVID GAUTHIER, MORALS BY AGREEMENT 165-89 (1986).

link with economics. For example, a prominent philosopher recently argued that the advantage a person gains from making a commitment provides a reason for carrying through later, even though the person subsequently can gain an advantage by not following through.[40] This argument connects to the economic literature on self-control. Unfortunately, each of the economic papers on self-control, including my own, seems like a fresh start rather than a cumulative contribution.[41]

A consensus on the foundations of moral commitment continues to elude economic theory. The concept of a Pareto self-

---

[41] See, e.g., THOMAS C. SCHELLING, CHOICE AND CONSEQUENCE (1984); THE MULTIPLE SELF (Jon Elster ed., 1986); SMOKING POLICY: LAW, POLITICS, AND CULTURE (Robert L. Rabin & Stephen D. Sugarman eds., 1993); Gary S. Becker et al., An Empirical Analysis of Cigarette Addiction, 84 AM. ECON. REV. 396 (1994); Gary S. Becker & Kevin M. Murphy, A Theory of Rational Addiction, 96 J. POL. ECON. 675 (1988); Frank Chaloupka, Rational Addictive Behavior and Cigarette Smoking, 99 J. POL. ECON. 722 (1991); Robert D. Cooter, Lapses, Conflict and Akrasia in Torts and Crimes: Towards an Economic Theory of the Will, 11 INT'L REV. L. & ECON. 149 (1991); Ted O'Donoghue & Matthew Rabin, Doing it Now or Later, 88 AM. ECON. REV. (forthcoming Sept. 1998); Rabin, Moral Preferences, supra note 34; Matthew Rabin, Cognitive Dissonance and Social Change, 23 J. ECON. BEHAV. & ORG. 177 (1994); Thomas C. Schelling, Addictive Drugs: The Cigarette Experiment, 255 SCIENCE 430 (1992); Ole-Jørgen Skog, The Strength of Weak Will, 9 RATIONALITY & SOC'Y 245 (1997); Thaler & Shefrin, supra note 32.

[42] Dixit and Norman observe that advertising changes preferences, so they evaluate the consequences of advertising from the viewpoint of initial preferences and final preferences. See Avinash Dixit & Victor Norman, Advertising and Welfare, 9 BELL J. ECON. 1 (1978); Avinash Dixit & Victor Norman, Advertising and Welfare: Reply, 10 BELL J. ECON. 728 (1979); Avinash Dixit & Victor Norman, Advertising and Welfare: Another Reply, 11 BELL J. ECON. 753 (1980). This approach resembles my own in this paper, except that I consider that an individual may choose whether to change his preferences.

improvement, which I introduce in this paper, seems quite novel, although some similarities exist with a series of papers by Dixit and Norman on advertising.[42]

Modern theories of the corporation emphasize the agency problem between owners and managers, or managers and workers.[43] Agency problems, which arise whenever the interests of the principals diverge from the interests of agents, impede cooperation and reduce its productivity. The agency problem in corporations expresses an old concern of sociologists in new language. Weber argued that the emergence of capitalism depended upon an ethic, first perfected among Protestant Christians, in which the individual internalized an occupational role.[44] "Internalization" here means accepting the norms of an occupation so intimately that they become part of the individual's self-conception, thus altering his perceived self-interest. Internalization of an occupational role, according to Weber, increases the dedication and creativity with which individuals pursue business goals. Similarly, Durkheim asked how a modern

---

[43] For modern views of the corporate form, see HENRY HANSMANN, THE OWNERSHIP OF ENTERPRISE (1996); Kenneth J. Arrow, The Economics of Agency, in PRINCIPALS AND AGENTS 37 (John W. Pratt & Richard J. Zeckhauser eds., 1985); Robert C. Clark, Agency Costs versus Fiduciary Duties, in PRINCIPALS AND AGENTS, supra, at 55.

[44] MAX WEBER, THE PROTESTANT ETHIC AND THE SPIRIT OF CAPITALISM 155-83 (Talcott Parsons trans., 1930) (1904-1905) (documenting the relationship between Protestant ascetism and the "spirit of capitalism").

[45] Durkheim frames his inquiry as follows:

society could divide labor so finely and still hold itself together.[45] He found the answer in the internalization of occupational roles.[46]

The application of science to industry requires decentralized decision making. According to Weber, certain Protestant groups internalized values emphasizing hard work, frugality, and honesty. The internalization of these values reduced the agency problem that plagues business. Casson revived this view in a recent book applying game theory to business practice.[47] Modeling business norms can give new vitality to an old vision of what makes capitalism possible.

II. SANCTIONS AS PRICES

I first view law externally, like the bad man of Holmes, and then I model self-control and self-improvement. Criminal codes order wrongs by seriousness: petty theft is less serious than grand theft, burglary is less serious than robbery, robbery is

---

[46] See id. at 345-50.

[47] Casson opens his book with the following observation:

> This book has a simple point to make. Overall economic performance depends on transaction costs, and these mainly reflect the level of trust in the economy. The level of trust depends in turn on culture. An effective culture has strong moral content. Morality can overcome problems that formal procedures--based on monitoring compliance with contracts--cannot. A strong culture therefore reduces transaction costs and enhances performance--the success of an economy depends on the quality of its culture.

MARK CASSON, THE ECONOMICS OF BUSINESS CULTURE 3 (1991).

less serious than assault, and so forth. As wrongdoing becomes more serious, the appropriate moral or legal sanction becomes more severe. The retributive theory of punishment formulates the relationship precisely: punishment should be proportional to the seriousness of the wrong.[48]

According to a conventional definition, a law is an obligation backed by a sanction.[49] The bad man views the sanction attached to an obligation as a price. The economic analysis of law built much of its early success on applying price theory to sanctions.[50] Sanctions, however, differ from market prices in important ways. Markets price permitted acts that leave the individual discretion to make choices. In contrast, law sanctions forbidden acts that involve social judgments. When committing a

---

[48] For a recent contribution to retributive theory, see Michael S. Moore, The Moral Worth of Retribution, in RESPONSIBILITY, CHARACTER, AND THE EMOTIONS 179 (Ferdinand Schoeman ed., 1987) (rebutting common arguments against retributive judgments).

[49] The view that law is the command of the sovereign or the obligations that the community imposes upon its members is so old that its origins cannot be determined. John Austin stated the imperative theory with clarity in 1 JOHN AUSTIN, The Province of Jurisprudence Determined, in LECTURES ON JURISPRUDENCE at 88-106 (photo. reprint, Thoemmes Press 1996) (Robert Campbell ed., 4th rev. ed. 1879). See also H.L.A. HART, THE CONCEPT OF LAW 18-25 (2d ed. 1994) (explaining the Austinian notion of "orders backed by threats"); JOSEPH RAZ, THE CONCEPT OF A LEGAL SYSTEM 5-26 (2d ed. 1980) (summarizing Austin's theory of a legal system).

[50] Price theory is not strategic, whereas game theory is strategic. Like the field of industrial organizations, the economic analysis of law first developed by heavy reliance on price theory and then gradually shifted to greater reliance on game theory. In other words, the economic models of law became increasingly strategic during the last decade.

forbidden act, the seriousness of the wrong depends upon the wrongdoer's attitude. The more disrespect the wrongdoer shows for the social judgment, the more serious is the wrong. To illustrate, wrongdoing becomes more serious as it passes from accidental to reckless, from reckless to intentional, from intentional to deliberate, and from deliberate to malicious.

These facts give sanctions a special character when viewed as prices. An obligation partitions acts into permitted and forbidden zones. When a penalty attaches to forbidden activities, private costs of the actor jump up at the point of the partition. Because of this discontinuity, most actors are not on the margin, so they do not respond to moderate changes in the magnitude of the sanction or the frequency of its exaction. They do respond, however, to changes in the partition's location. In brief, many actors respond to changes in legal obligations and relatively few actors respond to changes in the magnitude of the sanction or the frequency of its application.[51]

These facts distinguish a sanction attached to an obligation from a price attached to permission. Permission grants discretion to the individual to decide whether or not to do the act, so long as he pays the price. Prices do not usually create

---

[51] See Robert Cooter, Prices and Sanctions, 84 COLUM. L. REV. 1523, 1532 (1984) ("Most people conform to a reasonable obligation backed by a reasonable sanction, even if the legal standard is inefficient or otherwise undesirable. . . . [M]istakes in computing the level of the sanction or the

discontinuities in the costs faced by decision-makers. More actors, consequently, balance benefits and costs at the margin. When benefits and costs are equi-poised, a change in price will tip the balance and change behavior. Unlike a sanction, many individuals respond to changes in the magnitude of a price or the frequency of its exaction.

III. SELF-CONTROL

Having sketched the special kind of price created by attaching a sanction to an obligation, I now consider how a rational person might decide whether or not to commit a wrong. Assume that committing the wrong yields an immediate benefit at time 1, denoted b_1, and risks future punishment at time 2, denoted c_2 for cost. Let r denote the rate at which the actor discounts costs for futurity and uncertainty.[52] Thus the rational actor follows this rule:

$$b_1 - \frac{c_2}{r} \geq 0 \Rightarrow \text{commit the wrong;}$$

$$b_1 - \frac{c_2}{r} < 0 \Rightarrow \text{do not commit the wrong;}$$

The tipping point occurs where the actor is equi-poised between committing the wrong and not committing it. The tipping point value of r, denoted r^*, is found by solving the equation:

---

frequency of its application are not crucial, because most people will conform in spite of these mistakes.").

[52] The discount rate is sometimes written r and sometimes written $(1 + r)$. The point is that the value exceeds 1. To illustrate using the notation in this paper, the discount rate might be, say, $r = (1 + 0.07)$.

$$b_1 - \frac{c_2}{r^*} = 0,$$

which implies

$$r^* = \frac{c_2}{b_1}.$$

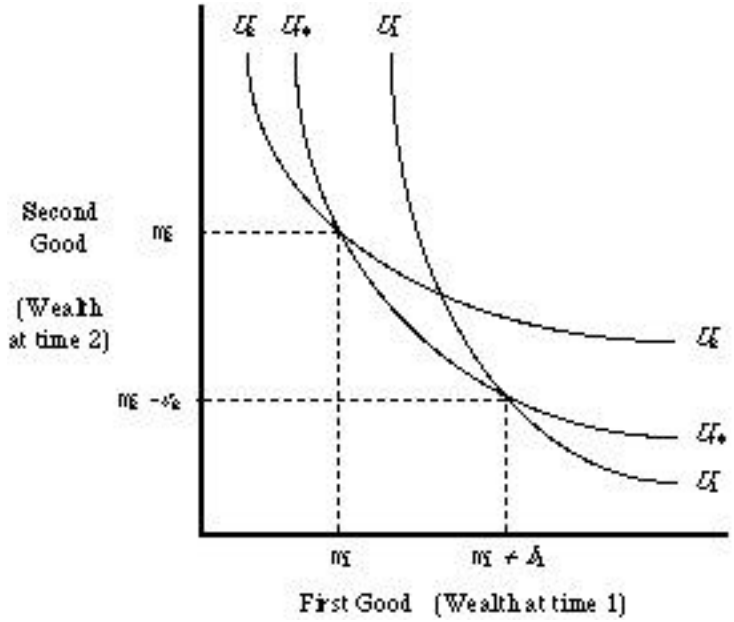
Thus an actor whose discount rate exceeds r* commits the wrong, and an actor whose discount rate falls short of r* does not commit the wrong.

A. Graphing Wrong and Right

Simplification of this decision problem permits its representation in Figure 1.[53] I will depict the fact that committing the wrong increases early income at the expense of later income. The horizontal axis in Figure 1 indicates wealth at time 1 and the vertical axis indicates wealth at time 2. If the actor does not commit the wrong, he enjoys wealth w_1 at time 1 and he enjoys wealth w_2 at time 2. Alternatively, committing the wrong pays b_1 at time 1, thus increasing his wealth to $w_1 + b_1$ at time 1. Furthermore, committing the wrong costs c_2 at time 2, thus decreasing his wealth to $w_2 - c_2$ at time 2.

[53] To simplify in order to use Figure 4, I assume that all relevant values are monetary and that future punishment of wrongdoing is certain.

FIGURE 1
TIPPING IN AND OUT



Now I add indifference curves to the graph in order to depict whether the actor prefers doing right or wrong. The actor's discount rate $\underline{r}$ determines the shape of the indifference curves in Figure 1. Notice that indifference curve $\underline{U}_{r^*}$ goes through ($\underline{w}_1$, $\underline{w}_2$) and ($\underline{w}_1 + \underline{b}_1$, $\underline{w}_2 - \underline{c}_2$). The actor with preferences $\underline{U}_{r^*}$, therefore, is indifferent between committing the wrong or not committing it. $\underline{U}_{r^*}$ is the indifference curve corresponding to the tipping value $\underline{r}^*$.

I also depict the indifference curves for an actor who prefers doing wrong, and another actor who prefers doing right: ($\underline{w}_1 + \underline{b}_1$, $\underline{w}_2 - \underline{c}_2$) lies on $\underline{U}_1$, and ($\underline{w}_1$, $\underline{w}_2$) lies below $\underline{U}_1$.[54] The actor with preferences $\underline{U}_1$, therefore, prefers to do wrong.

---

[54] In terms of the discount rate $\underline{r}$, $\underline{U}_1$ implies a discount rate $\underline{r}$ such that $\underline{r} > \underline{r}^*$.

Conversely, ($w_1 + b_1$, $w_2 - c_2$) lies below U_2, and (w_1, w_2) lies on U_2.[55] The actor with preferences U_2 therefore prefers to do right. The preferences U_{r^*} indicate the tipping value for r, and U_1 tips into committing the wrong, whereas U_2 tips out of committing it.

B. Variable Discounting

Now I turn to the problem of self-control. Like Milton's Satan, some bad people exercise self-control in pursuing bad ends. More often, however, lack of self-control contributes to wrongdoing. I will develop Figure 1 to show how to model the connection between lack of self-control and wrongdoing. As moods shift, a person may discount the future at different rates, so the same person's preferences may vary among U_{r^*}, U_1, and U_2 at different points in time. Given variation, a person prefers wrong when he has preferences U_1, and he prefers right when he has preferences U_2.

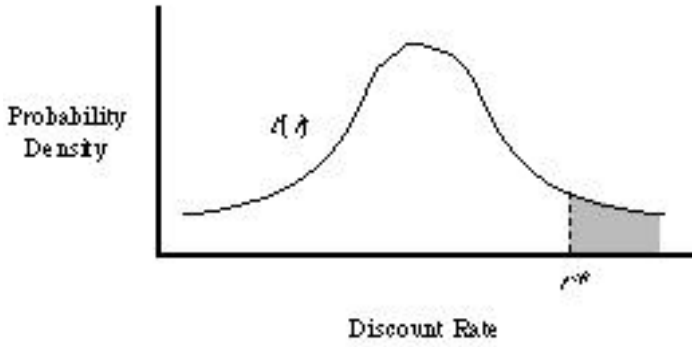
Generalizing, the discount rate r corresponds to different preferences and can vary continuously. Let Figure 2 depict a distribution $f(r)$ of values of r. If the actual value r drawn from the distribution $f(r)$ equals or exceeds r^*, the actor commits the wrong. The small shaded area in the right-tale of the distribution represents the probability that the actor commits the wrong. Conversely, if the actual value drawn from the distribution $f(r)$ is less than r^*, the actor does <u>not</u> commit the

---

[55] In terms of the discount rate r, U_2 implies a discount rate r such that $r < r^*$.

wrong. The unshaded area in the distribution represents the probability that the actor does not commit the wrong.

FIGURE 2
DISCOUNTING



A person's discount rate for uncertainty and futurity varies with mood, and changes in mood obey a mysterious chemistry. For purposes of my model, mood fluctuates randomly, causing the discount rate r to fluctuate randomly. Figure 2 represents the probability distribution from which the actor draws a discount rate whenever he makes a decision. With low probability, the actor draws a value of r greater than r* and commits the wrong, as with U_1 in Figure 1. With high probability, the actor draws a value of r smaller than r* and does not commit the wrong, as with U_2 in Figure 1.

Whenever the actor draws a discount rate close to the tipping value r*, a small change in punishment c can tip the decision one way or another. For example, a small increase in punishment causes the actor to decide against committing the wrong, whereas a small decrease in punishment causes the actor to decide in favor of committing the wrong. The probability that the

actor draws a discount rate close to $\underline{r}^*$ is low, whereas the probability that the actor draws a discount rate much smaller than $\underline{r}^*$ is high. When $\underline{r}$ is much smaller than $\underline{r}^*$, a small change in punishment cannot tip the decision one way or another.

An increase in the variability of moods increases the probability of wrongdoing by the actor. In terms of Figure 2, spreading the distribution by shifting density into the tails increases the area to the right of $\underline{r}^*$.[56] Greater probability density to the right of $\underline{r}^*$ implies an increase in the probability of wrongdoing. Conversely, a decrease in the variability of moods decreases the probability of wrongdoing.

Moods are more variable in youth than in old age. For spontaneous wrongdoing influenced by mood, the state should punish the young and old differently. Elsewhere I have shown that optimal deterrence requires young offenders to receive relatively mild punishment with high probability, and old offenders to receive relatively severe punishment with low probability.[57] Instead of exploring optimal punishments for deterrence, however,

---

[56] To be precise, the probability of wrongdoing may increase, and cannot decrease, with a mean-preserving spread in the distribution of the actor's subjective discount rate. By definition, a "mean-preserving spread" in any probability distribution shifts density from the center to the tails so that the mean remains constant and the variance increases. See Cooter, supra note 41 at 153 (demonstrating in Proposition 2 that a "mean-preserving spread in the distribution of preferences towards risk . . . may increase and cannot decrease the probability of a [temporary mistake in preferences]")

[57] See id. at 154-55 & 155 n.10 (observing that young people have a lower perception of risk of punishment than older people).

I want to focus upon how the law changes people.

In order to decrease the probability of wrongdoing, families, schools, and other institutions of socialization help young people achieve self-control by reducing variability in their moods. By regulating these institutions, law influences socialization. Instead of focusing upon such regulations, however, I want to focus on how punishment changes people. As mentioned above, a law is sometimes defined as an obligation backed by a sanction.[58] Indeed, the imperative theory of law regards state sanctions as law's essence.[59] By showing how sanctions change people, I will show that socialization is an essential, necessary effect of law.

IV. SELF-IMPROVEMENT

Having modeled self-control, I now turn to self-improvement, by which I mean changing your own preferences to improve them. I discuss how sanctions prompt a rational person to change his preferences. First I will extend the familiar concept of Pareto efficiency. Figure 3 depicts a Pareto improvement in resource allocation, which means a change that benefits someone without harming anyone. Assume an initial allocation of resources that produces $\underline{w}_1$ of the first public good and $\underline{w}_2$ of the second public

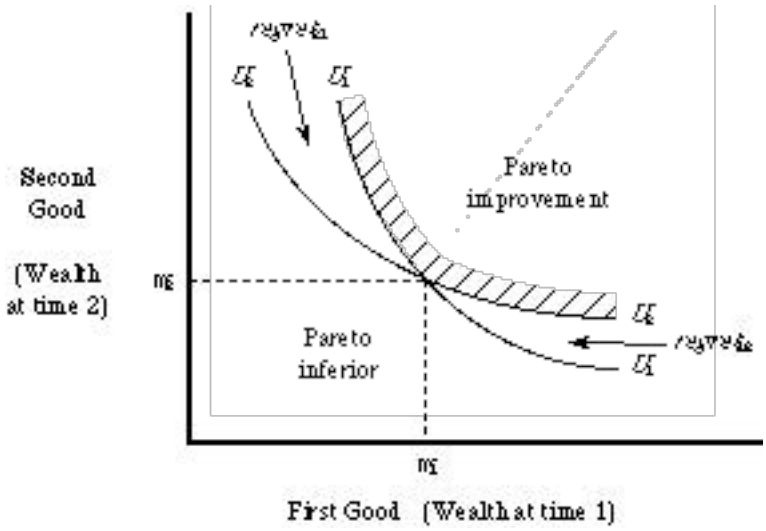
---

[58] See supra note 49 and accompanying text.

[59] See supra note 7 and accompanying text.

good.[60] This allocation enables person 1 to achieve utility $\underline{U}_1$ and person 2 to achieve utility $\underline{U}_2$. A Pareto improvement is a change that causes an increase in utility for at least one person without a decrease in the other's utility. The set of points indicated by hatch marks in Figure 3 contains the Pareto improvements relative to point ($\underline{w}_1$, $\underline{w}_2$).

FIGURE 3
PARETO IMPROVEMENT



I extend the idea of a Pareto improvement by developing the analogy between different people at the same time and the same person at different times. Reinterpret Figure 3 as depicting a single person with different preferences at different times. At

---

[60] For now, the two goods $\underline{w}_1$ and $\underline{w}_2$ should be interpreted as public. Public goods simplify the representation, because each person enjoys the same quantity of goods. A more conventional representation, which is unnecessarily complicated, uses private goods and an Edgeworth box. See **[FUDENBERG & TIROLE, supra note 34, at ___ (**parenthetical**)]**. Later I will intepret Figure 3 as depicting different preferences of the same person. See infra notes 61-63 and accompanying text. With this change, $\underline{w}_1$ and $\underline{w}_2$ in Figure 3 can be interpreted as private goods.

time 1 the actor in Figure 3 enjoys the allocation of goods ($\underline{w}_1$, $\underline{w}_2$) which yields utility $\underline{U}_1$.[61] At time 2 the actor's preferences change to $\underline{U}_2$.[62] The hatch marks in Figure 3 indicate the set of points above indifference curves $\underline{U}_1$ and $\underline{U}_2$. I describe the points in this set as Pareto improvements relative to point ($\underline{w}_1$, $\underline{w}_2$) for the same individual with different tastes. The analysis in Figure 3 easily generalizes from two preferences and two periods to many preferences and many periods.[63]

A. REGRET

With changing preferences, regret occurs when a choice produces a better result from the viewpoint of the initial preferences and a worse result from the viewpoint of final preferences. To illustrate, consider possible changes from the initial point ($\underline{w}_1$, $\underline{w}_2$) in Figure 3. The wedge between the utility curves, labeled $\underline{\text{regret}}_{12}$ in Figure 3, indicates points the actor would prefer with preferences $\underline{U}_1$ and regret with preferences $\underline{U}_2$. If preferences change from $\underline{U}_1$ at time 1 to $\underline{U}_2$ at time 2, then a decision by the actor at time 1 to choose a point in the set

---

[61] In this interpretation, the goods $\underline{w}_1$ and $\underline{w}_2$ can be public or private without altering the argument.

[62] The fact that $\underline{U}_1$ intersects $\underline{U}_2$ indicates that the actor's preferences have changed.

[63] Pareto improvements can be measured relative to any number of preferences, including the initial and final preferences of the affected party, and the preferences of the policy maker. To represent this generalization, add additional utility curves through point ($\underline{w}_1$, $\underline{w}_2$) in Figure 3 to indicate additional preferences, then draw the upper envelope to represent the set of Pareto improvements.

$\underline{regret}_{12}$ would cause regret at time 2.

In the absence of self-control, regret follows a predictable pattern, as depicted in Figure 1. In the predictable pattern, a youth with preferences $\underline{U}_1$ in Figure 1 highly discounts the future and commits a wrong, thus choosing ($\underline{w}_1 + \underline{b}_1$, $\underline{w}_2 - \underline{c}_2$). The youth matures into an adult with preferences $\underline{U}_2$ in Figure 1, who moderately discounts the future and prefers doing right to obtain ($\underline{w}_1$, $\underline{w}_2$). Thus the mature adult regrets doing wrong as a youth. With fluctuating preferences, an actor conceivably could get trapped in a cycle of regret.[64]

The possibility of regret conveys an advantage upon Pareto improvements. Since Pareto improvements are better from the viewpoint of the initial preferences and final preferences, the actor cannot regret a Pareto improvement. To illustrate, the set of points indicated by hatch marks and labeled "Pareto improvement" in Figure 3 does not intersect the set of points labeled $\underline{regret}_{12}$ or $\underline{regret}_{21}$. Consequently, no point causing regret is a Pareto improvement.

---

[64] To illustrate, on Monday morning, I might regret something that I did wrong on Saturday night. On Saturday night, however, I regret not having taken advantage of the opportunity that I had to do something wrong on Monday morning. To illustrate formally, the wedge between the utility curves, labeled $\underline{regret}_{21}$ in Figure 3, indicates points the actor would prefer with preferences $\underline{U}_2$ and regret with preferences $\underline{U}_1$. Assume the actor has preferences $\underline{U}_2$ at time 2, and then reverts to preferences $\underline{U}_1$ at time 3. A decision by the actor at time 2 to choose a point in the set $\underline{regret}_{21}$ would cause regret at time 3. An actor whose preferences fluctuate between $\underline{U}_1$ and $\underline{U}_2$ might make choices that fluctuate between $\underline{regret}_{12}$ and $\underline{regret}_{21}$, always regretting his choice.

B. When Preferences Influence Opportunities

Now I use the concept of a Pareto improvement to explain why a person might deliberately change his preferences. Good character makes someone more valuable in a cooperative venture. Participants in cooperative ventures often get paid according to their contribution. Good character can thus convey an advantage in cooperative ventures. For example, a person with more self-control may have more opportunity to work in jobs that demand reliability. Similarly, a more honest person may have more opportunities to manage valuable assets. Preferences influence opportunities, or, in plain speech, who we are influences what we can get.

To reward character, people must observe it in others. One person can observe another's character, although imperfectly. So, I say that character is translucent--not opaque and not transparent. Faking good character is an art that requires talent and skill, especially in enduring relationships.[65] In some circumstances, acquiring good character is the cheapest way to appear to have it. In this paper I will not analyze how people signal their character or observe it in others. Instead, I simplify by assuming that character is observable.

Exactly how people who want to improve their character succeed in doing so remains murky. Better character is not

---

[65] Joke--Sincerity is the key to success. Once you learn to fake it, you can do anything.

obtained merely by wanting it, although wanting better character may be necessary to obtaining it.[66] Presumably people improve their character by the same means that parents stress with their children, such as good habits, good associations, and moral or religious education. In this paper I will not analyze how people improve their character. Instead, I simplify by assuming that character is chosen.

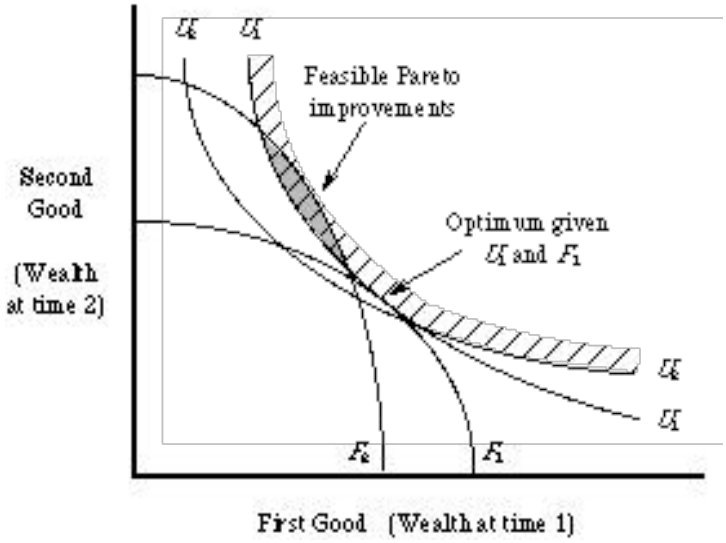
I will refer to the opportunities available to an actor as the "feasible set." Figure 4 depicts a feasible set $\underline{F}_1$ indicating the opportunities available to an actor with preferences $\underline{U}_1$.

Figure 4 also depicts a feasible set $\underline{F}_2$ indicating the alternatives available to an actor with preferences $\underline{U}_2$. Assume that an actor with preferences $\underline{U}_1$ can choose to retain the same preferences $\underline{U}_1$ and opportunities $\underline{F}_1$, or acquire new preferences $\underline{U}_2$ and opportunities $\underline{F}_2$. Would a rational actor make the change?

---

[66] Joke--How many psychiatrists does it take to change a light bulb? One, but the light bulb must want to change.

FIGURE 4
PARETO SELF-IMPROVEMENT



The standard of Pareto improvements provides a compelling answer. Given preferences $\underline{U}_1$ and feasible set $\underline{F}_1$, the actor's initial optimum occurs at the point where $\underline{F}_1$ is tangent to $\underline{U}_1$, as indicated in Figure 4. As in Figure 3, the hatch marks in Figure 4 indicate the set of Pareto improvements relative to the initial optimum. Some of the Pareto improvements are feasible with opportunities $\underline{F}_2$. Specifically, the shaded lozenge contains the <u>feasible</u> Pareto improvements. Thus the actor who changes preferences from $\underline{U}_1$ to $\underline{U}_2$ creates the opportunity for a better payoff as measured by original preferences or final preferences. So the actor has a strong reason to make the change.

In general, I will use the phrase "Pareto self-improvement" to mean a change made by the actor in his preferences that makes feasible an allocation preferred by original preferences and

final preferences.[67]

C. Example: Work Ethic as Pareto Self-improvement

Now I relate my model of self-improvement to an aspect of the work ethic that Weber attributed to Protestantism.[68] Consider a worker who can choose whether or not join a religious sect and internalize a work ethic. The work ethic values production and devalues leisure. An employer rationally expects a convert to such a sect to work more and relax less, which makes the worker more valuable to the employer. Thus internalizing this ethic can increase opportunities to earn income from work.

I can reinterpret Figure 4 to fit this example. The first good on the horizontal axis can be interpreted as leisure, and the second good on the vertical axis can be interpreted as income. Thus a person with preferences U_1 likes leisure, whereas the person with preferences U_2 internalizes the work ethic and likes income. The person who likes work has more opportunities for income and fewer opportunities for leisure, as indicated by F_2, relative to the person who likes leisure and has

---

[67] Notice that changing preferences must cause a Pareto improved allocation to become feasible. I do not require that the allocation actually chosen with the change in preferences be a Pareto improvement. I could mark the difference by distinguishing hypothetical Pareto self-improvements (in which an actual Pareto improvement is feasible) and actual Pareto self-improvements (in which a Pareto improvement actually is made). The difference could be significant for some kinds of moral problems. At this early stage of developing the theory, however, I dispense with refinements.

[68] See supra note 44 and accompanying text.

opportunities F_1. In Figure 4, internalizing the work ethic is a Pareto self-improvement, so the actor has a strong reason to do it.

D. How Law Prompts Self-improvement

I will explain some ways that law can change preferences. An employer, partner, lender, friend, or spouse cares about the character of the other party in the relationship. In private relationships, character gets rewarded or punished. In tribal law, anthropologists observe that dispute resolution focuses on relationships more than acts, and relationships deeply implicate character.[69] In modern law, courts sometimes modulate sanctions according to the actor's character, as with a disloyal fiduciary or a vicious criminal. In my discussion of Pareto self-improvement, I mentioned that acquiring good character is sometimes the cheapest way to appear to have it.[70] So sanctioning apparent character can cause people to change their actual character.

More typically, however, modern courts sanction acts without inquiring into character, even though character may change in response. In general, the law prompts changes in character whenever a legal sanction creates an opportunity for Pareto self-

---

[69] See, e.g., Paul Bohannon, Justice and Judgment Among the Tiv 33-37 (2d prtg. 1968) (observing that in Tiv judicial proceedings, witnesses must possess a close relationship with an accused, and that the relationship may affect the perceived truthfulness of the accused).

[70] See supra Part IV.B.

improvement.

For example, law can discourage work by taxing income from labor. A heavy tax on income from labor reduces the appeal of internalizing the work ethic, as I explain using Figure 4. Recall that the person who internalizes the work ethic changes preferences from $\underline{U}_1$ to $\underline{U}_2$, and the resulting change in opportunities from $\underline{F}_1$ to $\underline{F}_2$ is a Pareto self-improvement. A heavy tax would reduce the after-tax income obtainable from work, causing the opportunity sets in Figure 4 to shift down. After imposing a heavy tax on income in Figure 4, internalizing the work ethic may no longer be a Pareto self-improvement.

As another example, consider any situation where society informally punishes bad character and the state formally punishes the associated wrongdoing. Examples include a fiduciary's disloyalty and diversion, a promisor's dishonesty and breach of promise, and borrower's recklessness and misuse of funds. Adding formal punishment by the state to the informal punishment by society may be enough to make acquiring good character a Pareto self-improvement.

I will illustrate these facts using a variant of the model of time-discounting. Recall that doing right, which involves cooperating with others, yields a relatively low payoff in the first period and a relatively high payoff in the second period.[71] The total subjective payoff that an individual receives from

cooperating depends upon his discount rate. A person with a low discount rate, who appears to be responsible, receives a relatively important job that yields the relatively high payoff denoted ($\underline{w}_1$, $\underline{w}_2$). In contrast, a person with a high discount rate, who appears to be irresponsible, receives a relatively unimportant job that yields a relatively low payoff.[72]

Instead of doing right and cooperating, a person can do wrong and not cooperate. Wrongdoing yields a relatively high payoff $\underline{w}_1 + \underline{b}$ in the first period and a relatively low payoff $\underline{w}_2 - \underline{c}$ in the second period. To keep the analysis simple, I assume that wrongdoers, who do not remain in a cooperative venture for long, receive the same payoff regardless of their discount rate.

To show the role of law explicitly, I decompose the cost $\underline{c}$ into two elements. The first element is a social sanction $\underline{c}_s$, and the second element is a legal sanction $\underline{c}_\lambda$, where the total sanction $\underline{c}$ is equal to $\underline{c}_s + \underline{c}_\lambda$.

Figures 5A and 5B depict the payoffs and utilities for the two types of people. In Figure 5A, a person with high discount rate $\underline{U}_1$ would prefer to have the payoff ($\underline{w}_1$, $\underline{w}_2$) available to a person with the low discount rate who does right, rather than doing wrong and receiving the payoff ($\underline{w}_1 + \underline{b}$, $\underline{w}_2 - \underline{c}_s - \underline{c}_\lambda$).

---

[71] See supra note 54 and accompanying text.

[72] I have no need to represent the low payoff in notation.

Similarly, in Figure 5B, a person with low discount rate $\underline{U}_2$ prefers to do right and receive ($\underline{w}_1$, $\underline{w}_2$), rather than doing wrong and receiving the payoff ($\underline{w}_1$ + $\underline{b}$, $\underline{w}_2$ - $\underline{c}_s$ - $\underline{c}_\lambda$). I have shown that acquiring the low discount rate $\underline{U}_2$ is a Pareto self-improvement for the actor.

FIGURE 5
STATE SANCTIONS AND SOCIAL SANCTIONS

FIGURE 5A

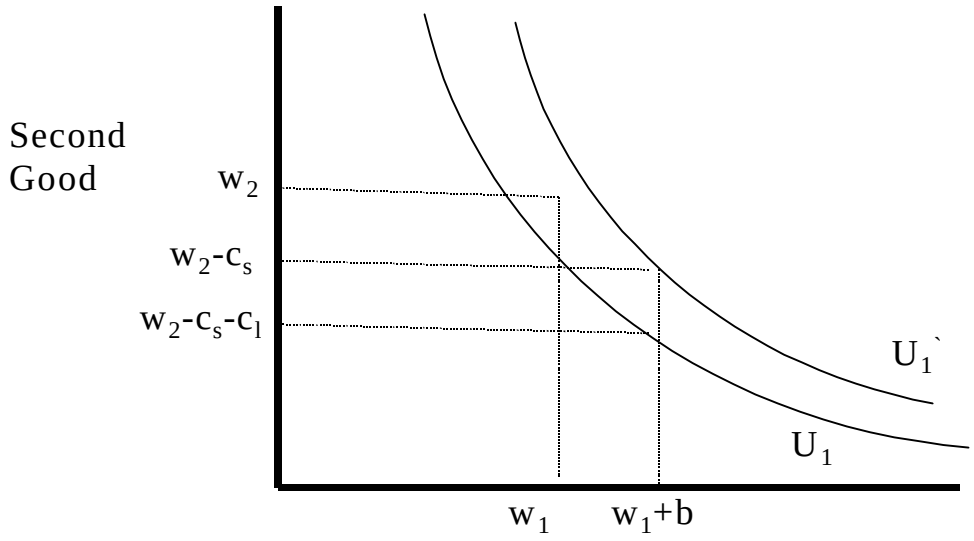
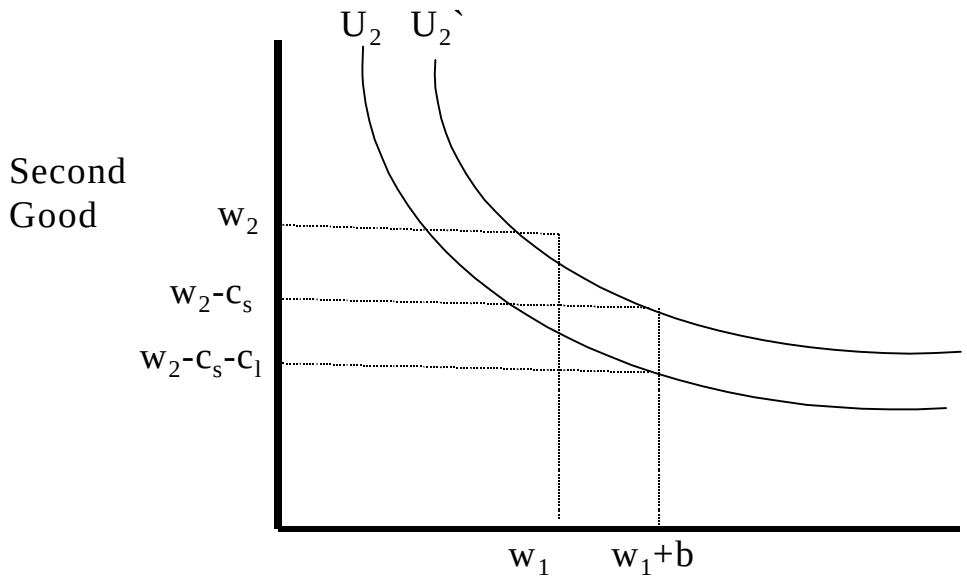


FIGURE 5B



This demonstration assumed a social sanction c_s and a legal sanction c_λ. Now consider the consequences of removing the legal sanction c_λ. Without any legal sanction, the wrongdoer would only suffer the social sanction c_s. A person with low discount rate U_2 in Figure 5B prefers (w_1, w_2) rather than (w_1 + b, w_2 - c_s). So, after removing the legal sanction, the person with low discount rate continues to do right. In contrast, a person with high discount rate U_1 in Figure 5A prefers (w_1 + b, w_2 - c_s) rather than (w_1, w_2). So, after removing the legal sanction, the person with high discount rate prefers to do wrong rather than acquire a low discount rate and do right. Without a legal sanction for wrongdoing in Figure 5A and Figure 5B, low discounting is not a

Pareto improvement over high discounting.

Besides sanctions, the state might resort to more manipulative policies to change preferences. To illustrate, assume that instead of imposing money sanctions, the state could shame wrongdoers by publicizing their misdeeds.[73] Like other sanctions, the prospect of shaming could tip the balance in favor of improving one's character. Replacing liability with shaming for a particular class of wrongdoing might make some people better off relative to their initial and final preferences without making anyone worse off. Under these assumptions, everyone, including cynics, might agree to replace liability with shaming as the sanction for the wrongdoing in question.

E. Advantages of the Concept of Pareto Self-improvement

Why adopt this novel concept? For the same reasons that economists adopted the original concept, which has a positive and a normative use. When a situation is Pareto inefficient, people recognize that a change could benefit someone without harming anyone. These facts create pressure for change without counter-pressure to resist change. Consequently, Pareto efficient situations tend to be more stable than Pareto inefficient situations. This use of Pareto efficiency is predictive.

Pareto efficiency with variable tastes should have similar

---

[73] The possible revival of shaming is discussed in Dan M. Kahan, Social Influence, Social Meaning, and Deterrence, 83 Va. L. Rev. 349, 384-85 (1997) (extolling the use of shaming as a worthwhile alternative sanction that costs less than incarceration).

stability properties as Pareto efficiency with fixed tastes. People with opportunities for Pareto self-improvements will change themselves. After exhausting all Pareto self-improvements, any further changes in character encounter psychological resistance. An individual might regret further changes.

The decision to make a Pareto self-improvement does not require the individual to compare one set of preferences to another. For example, the individual does not have to decide whether being reliable is intrinsically better than being unreliable. Similarly, the individual in Figure 4 has a reason to act without knowing whether preferences U_2 are inherently better or worse than preferences U_1. In this respect, the individual does not need a deep ethical theory.[74] The individual who lacks a deep ethical theory, or the individual whose ethical theory does not apply to the choice in question, or the individual who does not have the time and inclination to figure out how his ethical theory applies to the choice in question, can still apply the Pareto criterion.

In contrast, choosing among Pareto efficient points by changing character requires a deep ethical theory and much information. To illustrate, assume that a person in a certain job must choose between honesty with low profits and dishonesty with high profits. To make the choice, the person must have an ethical

---

[74] In this context, a "deep theory" corresponds to a cardinal utility theory, which assigns weight to different preferences.

theory that compares the value of honesty to its cost. Specifically, the ethical theory must say whether the intrinsic value of honesty exceeds its material disadvantage. Many people cannot decide such questions without soul-searching or agony.

In addition to its use in individual decision making, intra-personal Pareto efficiency has a use in policy-making. When preferences change, some ethical theories prefer the original preferences while other ethical theories prefer the final preferences. This fact creates a dilemma for evaluating public policies that change preferences. The Paretian approach partly resolves the dilemma by allowing the state to create opportunities for individuals to make Pareto self-improvements that do not harm anyone else.

Ethical theories that respect individual autonomy should recognize the desirability of intra-personal Pareto efficiency, which is achieved when no one can make further Pareto self-improvements without harming someone else. Ethical theories that disagree about the best goals for people to pursue, or disagree about how to resolve conflicts among the goals of different people, might yet agree that people should be free to improve themselves to the maximum extent without harming others. Identifying intra-personal Pareto improvements thus reduces the extent of disagreement among people with different values.

Different ethical theories embrace different ideals of distribution. Many economists resist committing their subject to

a particular ethical theory or distributive ideal. Many different ethical theories acknowledge the value of Pareto efficiency. For example, materialists who favor maximizing social wealth, utilitarians who favor maximizing the sum of utilities, and Rawlsians who favor maximizing the well-being of the least advantaged person acknowledge that their ideal world is Pareto efficient.[75] Given this fact, economists use Paretian analysis to say something about public policy without taking sides in disputes about distribution. Pareto self-improvements could be put to the same use to evaluate policies that change preferences.

CONCLUSION

In the essay celebrated by this conference, Holmes admires value-autonomy, by which I mean individual values not shaped by law. The bad man who disobeys law when it serves his advantage has value-autonomy. The good man who disobeys law when it makes immoral demands has value-autonomy. I suspect, however, that most people are not so bad or so good. Rather, I suspect that most people internalize important values from law. Their character is value-dependent with respect to law. Perhaps the law's coercion of the bad man is the small, visible part of the iceberg, and law's improvement of ordinary people is the large, invisible part of the iceberg. If so, the bad-man theory of law is woefully

---

[75] Wealth and utility maximization obviously require Pareto efficiency. As for the maximin, Rawls asserts that this objective is consistent with Pareto efficiency. See RAWLS, supra note 39 at

incomplete.

Moralists have long understood that sanctions for wrongdoing create incentives for improving oneself, but this idea has eluded economic models. In the examples that I have developed, sanctions deter wrongdoing and improve people. Law can strengthen a moral consensus by tracking morality, or law can undermine morality by departing from it.[76] Laws that seem unjust or morally irrelevant do not breed respect. In special circumstances, instead of strengthening morality, law can crowd it out.[77] My formulation of

---

78-80 (explaining that "the difference principle is compatible with the principle of efficiency").

[76] A more difficult question concerns whether law can create a moral consensus where none exists. This is a question of the expressive power of the law, which I explore in Robert Cooter, Expressive Law and Economics, 27 J. LEGAL STUD. (forthcoming June 1998).

[77] Crowding out of morality by law is a special concern of Bruno Frey. See, e.g., **[Bruno S. Frey, NOT JUST FOR THE MONEY __-__ (1997) (**parenthetical**)]**; **[Bruno S. Frey, A Constitution for Knaves Crowds Out Civic Virtues, 107 ECON. J. 1043, __-__ (____, 1994) (**parenthetical**)]**; **[Bruno S. Frey et al., The Old Lady Visits Your Backyard: A Tale of Morals and Markets, 104 J. POL. ECON. ___, 1300-01 (1996) (observing that hazardous waste facility siting plans tend to crowd out public spirit, and arguing that increased compensation could counteract the crowding out effect)]**.

In a related phenomenon, competitive markets can reduce the reward for virtue by reducing the need for enduring relationships, in which case small, imperfect markets promote virtue and large, competitive undermine virtue. In Brennan and Hamlin's attractive phrase, competition "economizes on virtue." See Geoffrey Brennan & Alan Hamlin, Economizing on Virtue, 6 CONST. POL. ECON. 35, 54-55 (exploring the conceptions of "virtue" and "economizing," and concluding that "institutions that attempt to economize on virtue must be designed with care").

Pareto self-improvement brings this idea under the analytical power of economic models. More generally, the concept of Pareto self-improvement extends economic reasoning to endogenous preferences and the internalization of norms.