Do Good Citizens Need Good Laws? Economics and the Expressive Function

Michael D. Gilbert and Andrew T. Hayashi***

We explore how adding prosocial preferences to the canonical precaution model of accidents changes either the efficient damages rule or the harm from accidents. For a utilitarian lawmaker, making the potential injurer sympathetic to the victim of harm has no effect on either outcome. On the other hand, making injurers averse to harming others reduces the harm from accidents but has no effect on efficient damages. For an atomistic lawmaker — one who excludes prosocial preferences from social welfare — cultivating a taste for either harm aversion or perfect sympathy can reduce efficient damages, though neither has any effect on the amount of harm from accidents. On the other hand, causing people to act as if they are averse to harm creation, such as out of habit or moral obligation, reduces both the efficient amount of damages and total harm. In general, encouraging either a distaste for, or moral commitment against, harm creation is useful while inculcating sympathy for victims of harm is not.

---

* Martha Lubin Karsh and Bruce A. Karsh Bicentennial Professor of Law, University of Virginia.

** Class of 1948 Professor of Scholarly Research in Law, University of Virginia. This Article was prepared for a conference in honor of Bob Cooter titled "How Law Changes What You Want: Positive and Normative Effects of Law on Values and Preferences," Berkeley Law School, February 2020. We thank Jennifer Arlen, Jonah Gelbach, Alon Klement, Saul Levmore, Mike Livermore, Greg Mitchell, Ariel Porat, and other conference participants for helpful comments. We thank Thomas Langstaff for excellent research assistance.

Introduction

Do good citizens need good laws? One might suppose the answer is no.[1] If citizens lacked "tastes" like animus, envy, and greed (antisocial preferences), and if they possessed tastes like sympathy, compassion, and charity (prosocial preferences), then much law would seem superfluous. Good citizens do not commit murder, drive drunk, or steal jewelry, so the state should not need to ban those activities. Good preferences should tend to promote cooperation,[2] lower enforcement costs,[3] increase social welfare,[4] and perhaps eliminate the

---

1 Laws that coordinate behavior are an exception. Even good citizens might struggle to coordinate on matters like traffic (who yields?), and law can help. *See generally*, Richard H. McAdams, The Expressive Powers of Law: Theories and Limits (2015).

2 *See, e.g.*, Robert D. Cooter, *Decentralized Law for a Complex Economy: The Structural Approach to Adjudicating the New Law Merchant*, 144 U. Pa. L. Rev. 1643, 1675 (1996) ("[P]eople who internalize a business norm cause more cooperation in a community […]"); Robert Axelrod, *An Evolutionary Approach to Norms*, 80 Am. Pol. Sci. Rev. 1095 (1986) (explaining the evolution of cooperative norms, including norms internalized as preferences); Edna Ullmann-Margalit, The Emergence of Norms (1977) (arguing that good norms, which might be internalized as preferences, promote collective action); David P. Gauthier, *Morality and Advantage*, 76 Phil. Rev. 460 (1967) (arguing that good morals promote collective action).

3 *See, e.g.*, Richard A. Posner, *Social Norms and the Law: An Economic Approach*, 87 Am. Econ. Rev. 365, 366 (1997) ("The internalization of norms through habituation may seem highly efficient because it reduces the cost of compliance."); Gary S. Becker, Accounting for Tastes 225 (1996) ("Honesty, to take one example, enormously reduces the need to spend resources . . . on the protection of property […]"); Dennis Chong, Collective Action and the Civil Rights Movement 69 (1991) ("[T]he well-trained conscience is the best policeman.").

4 *See, e.g.*, Louis Kaplow & Steven Shavell, *Fairness Versus Welfare*, 114 Harv. L. Rev. 961, 1349 (2001) ("[A]dopting policies contrary to the current preferences of some individuals may change those preferences, and, over the long run, social welfare may rise as a result."); *id.* at 1349 n.941 ("[C]ultivating stronger positive other-regarding preferences . . . tends to be socially valuable […]"); Richard H. McAdams, *Relative Preferences*, 102 Yale L.J. 1, 80 (1992) ("Society is probably still better off with the norms against envy and relative gratification than it would be without them."); Cass R. Sunstein, After the Rights Revolution: Reconceiving the Regulation State 67 (1990) (describing scenarios in which changing preferences could promote welfare); Alfred Marshall, Principles of Economics 181-82 (1890) ("It would be a gain if the moral sentiment of the community could induce people to avoid all sorts of display of individual wealth.").

need for most law altogether. As Madison wrote, "[i]f men were angels, no government would be necessary."[5]

Given the apparent benefits, governments might like to "improve" citizens' preferences, such as by encouraging them to care about the well-being of their fellow citizens or the harm that their activities impose on others. How to do so? Education is one option, but law may be another. According to Aristotle, successful laws "make the citizen good by inculcating habits in them."[6] An economist might interpret Aristotle's claim as a statement about law's capacity to change people's preferences for the better. Many scholars have made this argument[7] and some evidence supports it. In the United States, for example, antidiscrimination laws apparently changed attitudes towards some disadvantaged groups.[8]

---

5 James Madison, The Federalist No. 51, 264 (Ian Shapiro ed., 2009).

6 Aristotle, Nicomachean Ethics 34 (Martin Ostwald trans., 1962) ("Lawgivers make the citizen good by inculcating habits in them, and this is the aim of every lawgiver; if he does not succeed in doing that, his legislation is a failure.").

7 For some recent examples, see Robert Cooter, *Expressive Law and Economics*, 27 J. Legal Stud. 585, 603-06 (1998) [hereinafter Cooter, *Expressive*] (arguing that law can change preferences); Robert Cooter, *Models of Morality in Law and Economics: Self-Control and Self-Improvement for the "Bad Man" of Holmes*, 78 B.U. L. Rev. 903, 924-27 (1998) (same); Kenneth G. Dau-Schmidt, *Legal Prohibitions as More Than Prices: The Economic Analysis of Preference Shaping Policies in the Law*, *in* Law and Economics: New and Critical Perspectives 153, 158 (Robin Paul Malloy & Christopher K. Braun eds., 1995) (discussing "a variety of laws that seem consciously aimed at influencing people's preferences"); Kenneth G. Dau-Schmidt, *An Economic Analysis of the Criminal Law as a Preference-Shaping Policy*, 1990 Duke L.J. 1 (1990) (arguing that criminal law aims to change preferences); Albert O. Hirschman, Rival Views of Market Society: And Other Recent Essays 146 (1986) ("A principal purpose of publicly-proclaimed laws and regulations is to . . . influence citizens' values and behavior codes."); Cass R. Sunstein, *Legal Interference with Private Preferences*, 53 U. Chi. L. Rev. 1129, 1137 (1986) [hereinafter Sunstein, *Interference*] (suggesting that law's "moral function" is to change "objectionable preferences"); Guido Calabresi, Ideals, Beliefs, Attitudes, and the Law: Private Law Perspectives on a Public Law Problem 84 (1985) ("Law . . . is fundamentally concerned with shaping tastes."); Richard B. Stewart, *Regulation in a Liberal State: The Role of Non-Commodity Values*, 92 Yale L.J. 1537, 1538 (1983) (arguing that nurturing certain values "is, and should be, an important objective of regulatory and administrative law").

8 *See, e.g.*, John J. Donohue, *Prohibiting Sex Discrimination in the Workplace: An Economic Perspective*, 56 U. Chi. L. Rev. 1337, 1338-339 (1989) (suggesting that Title VII changed attitudes towards women).

Why does law change preferences? How do tastes evolve as preference-shaping laws interact with other social and economic pressures?[9] Though important, we eschew these positive inquiries and focus on the normative question: *should* law change preferences? Notwithstanding Madison's optimism, the answer is not so simple. Like others, we value individual autonomy and worry about paternalism. A program to improve preferences might be indistinguishable from brainwashing.[10] But even if this concern is set aside, improving preferences does not translate straightforwardly into social benefits.

To show this, we study a simple model of precaution under strict liability involving a pedestrian, a driver, and a lawmaker who can inculcate prosocial preferences in the driver. For those prosocial preferences to make a difference, either by increasing the driver's precaution or reducing optimal damages, we must accept one of two contestable claims. We must assume (for economics cannot prove it for us) that the benefits of more precaution exceed the costs that an accident imposes on the driver's guilty conscience. Or we must embrace what we call the *virtue paradox*. The lawmaker must inculcate prosocial preferences in the driver, which leads to psychological harm, but ignore that psychological harm when calculating social welfare.

Despite these problems, we show that law could be a helpful teacher. It could especially help by encouraging aversion to causing harm rather than complete sympathy with those who suffer. Law could also help by encouraging the adoption of habits or moral commitments that cause people to act *as if* they had prosocial preferences. In fact, the only way to achieve both a reduction in the amount of efficient damages and the harm from accidents is by causing people to behave as if they were harm averse.

In Part I, we describe the fundamentally political and philosophical nature of the social welfare function and the decision to shape individual preferences. In Parts II and III, we study the effect of changing individuals' preferences given

---

9 For theories on how preferences change, see, for example, Cooter, *Expressive*, *supra* note 7, at 598-606 (describing "Pareto self-improvements"); Becker, *supra* note 3, at 3-23 (connecting "personal" and "social" capital to preference change); Jon Elster, Sour Grapes: Studies in the Subversion of Rationality 109-33 (1983) (analyzing "adaptive" preferences). For a review of the scholarship on endogenous preferences, see Robin Hahnel & Michael Albert, Quiet Revolution in Welfare Economics 76-109 (1990). *See also* Henry J. Aaron, *Distinguished Lecture on Economics in Government: Public Policy, Values, and Consciousness*, 8 J. Econ. Persp. 3 (1994) (arguing that economists should focus on preference formation); Robert Cooter, *Law and Unified Social Theory*, 22 J.L. & Soc'y 50 (1995) (same).

10 Sunstein calls this the "liberty" objection to preference change. *See* Sunstein, *Interference*, *supra* note 7, at 1131-132.

two different social welfare functions: "utilitarian" and "atomistic." Part IV explores the benefits of using law to shape individuals' moral commitments, rather than their preferences.

I. Choosing Welfare

To claim that any action — saving a kitten, toppling a dictator — increases social welfare requires being explicit about how social welfare is calculated. One view, which dominates in economics, is that social welfare equals the total utility of all individuals in society, where utility is an ordinal numerical measure that represents an individual's preferences.[11] This approach works, at least conceptually, when preferences stay fixed, but it stumbles when preferences change.[12]

Take a simple example: Adam likes to run. In the winter, cold weather and slippery roads make it practically impossible to run. For Adam, running is a preference, and the climate is a constraint. Adam's utility will increase in the spring when the climate changes. Assessing Adam's utility is simple given changing constraints. But now change preferences. Today Adam prefers reading, but tomorrow he will prefer running. Will his utility increase tomorrow? Is Adam the Runner better off when running than Adam the Reader is when reading? Both Adams satisfy their preferences, so preference satisfaction, by itself, is not enough to answer the question. We must assess the preferences themselves.

The challenge increases in scale when we consider social welfare. If a change in tastes has an indeterminate effect on Adam, it must have an indeterminate effect on society.[13] If many people's tastes change, we cannot

---

11 On other approaches to social welfare, see, for example, Matthew D. Adler, Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis (2012).

12 We are certainly not the first to make this point. *See, e.g.*, Cooter, *Expressive*, *supra* note 7, at 602 (recognizing the challenge of assessing public policy and social welfare given changing preferences); Becker, *supra* note 3, at 21 (same); Hahnel & Albert, *supra* note 9, at 145-84 (same); Elster, *supra* note 9, at 134-36 (same); *cf.* Avinash Dixit & Victor Norman, *Advertising and Welfare*, 9 Bell J. Econ. 1 (1978) ("Welfare analysis of advertising usually begins and ends with the remark that there is no fixed standard for value judgment when tastes are variable.").

13 David Dolinko, *The Perils of Welfare Economics*, 97 Nw. U. L. Rev. 351, 383 (2002) ("[I]f something is conceptually awry with the comparison of one single individual's well-being before and after her preferences change, the same problem

assess welfare without making many choices about which preferences are better or worse, and for whom.

To illustrate this problem in the legal context, and to lay some groundwork, we revisit a textbook model of precaution.[14] Consider a driver going from home to work. She can take precautions (driving more slowly, stopping fully at stop signs) at cost x. The more precautions she takes, the higher her costs, so x increases. Precaution reduces the probability $p(x)$ of an accident.[15] In the event of an accident, a pedestrian suffers monetary harm h. The driver is subject to strict liability for any harm that she causes, in which case she pays the pedestrian damages equal to D. The driver's payoff depends on her expected costs, so we can express her utility function as a cost function: $x + p(x)D$. The driver would like to minimize that cost function. By the same logic, the pedestrian's cost function is $p(x)(h - D)$, and social costs are the sum of the two: $x + p(x)h$.

The driver in our scenario is selfish. She does not care about the well-being of the pedestrian; she cares about an accident only to the extent that she must pay damages. Suppose we can change the driver's preferences so that she dislikes imposing harm on others. Specifically, the driver's cost function becomes $x + p(x)(D + \alpha h)$, where $\alpha > 0$ indicates the weight the driver attaches to any harm she causes. The social cost function is now $x + p(x)(h + \alpha h)$.

Is society better off when the driver dislikes imposing harm? For economists, the question would seem unanswerable. Changing the driver's preferences changes the social cost function from $x + p(x)h$ to $x + p(x)(h + \alpha h)$. Although the second cost function appears larger, the costs captured by the two functions are denominated in units unique to each. The two functions are like different economies, each with its own currency, and with no market for exchanging them. Economics can tell us how to minimize each function, but it cannot tell us which function is *better*. Solving the models — predicting in each case the driver's behavior and the probability of an accident — would not help. If we do not know which function is better, we cannot determine whether the equilibrium under one, whatever it happens to be, is superior to the equilibrium under the other.

---

will infect any evaluation of social well-being before and after individuals' preferences change.").

14 Specifically, we study a textbook model of efficient, unilateral precaution. *See, e.g.*, Steven Shavell, *Strict Liability Versus Negligence*, 9 J. Legal Stud. 1, 2-6, 10-17 (1980).

15 We assume that the continuous and differential probability function has properties $p'(x) < 0 < p''(x)$.

Not everyone is so hamstrung. Some scholars have taken positions on how to choose between competing social welfare functions. Some positions are sophisticated,[16] while others are not.[17] We need not engage in that debate to make our point. In general, economics cannot tell us if any social welfare function is better than another. Assessing preferences is an exercise in philosophy, not economics. This is not news to most economists, but it might surprise other scholars, including legal academics who draw on economic reasoning in their work.

For these reasons, it is unconventional for economists to think of changing people's preferences as a tool for achieving a policy objective. Satisfying individual preferences, whatever they are, usually *is* the policy objective. Thus, we must frame the problem differently. The spirit of our inquiry is to evaluate some concrete consequences of prosocial preferences, to provide clarity to lawmakers evaluating the desirability of changing the things people value. In "The Future of Law & Economics," Professor Calabresi defends just this role for law and economics scholars, and ours is a modest first step in this research agenda.[18]

We imagine a society composed of people who are selfish and people who care about others. The lawmaker for this society cares about maximizing social welfare and will choose the legal rules that do so. However, the lawmaker also cares about ensuring that legal rules are enforceable and about minimizing the amount of harm individuals impose on others.

The first question the lawmaker must answer is whether the social welfare function should be *utilitarian*, meaning that it is simply the sum of all individual

---

16 *See, e.g.*, Burton A. Weisbrod, *Comparing Utility Functions in Efficiency Terms or, What Kind of Utility Functions Do We Want?*, 67 Am. Econ. Rev. 991, 994 (1977) ("The customary proposition that one type of utility function cannot be compared to another within an economic efficiency framework is correct in general. This note has suggested, however, that despite the general noncomparability of utility functions, some can be compared and found to be preferred to others."). *Cf.* Cooter, *Expressive*, *supra* note 7, at 602 (proposing a "Pareto criterion" that avoids the dilemma of prioritizing pre-change and post-change preferences).

17 *See* Kaplow & Shavell, *supra* note 4, at 1337 n.917 ("Often analysts simply suppose that a change in preferences would be desirable. That certain preferences are good and others are bad may seem obvious within some community of discourse (say, among a small group of philosophers), but its members may differ in relevant respects from most individuals in society . . . It is easy, but dangerous, for such a group simply to impose its notion of the good life on others […]").

18 Guido Calabresi, The Future of Law and Economics: Essays in Reform and Recollection (2016).

utilities, or whether it should be *atomistic*, meaning that the social welfare function disregards any psychic benefits or costs that flow from prosocial preferences but is otherwise utilitarian.[19] This is a fundamental, normative choice that must be decided before anything else, and which we do not try to resolve here. The second question the lawmaker faces is whether to change individuals' preferences to make them prosocial. It is straightforward to determine the legal rule that maximizes social welfare after the lawmaker has decided on individuals' preferences and the form of the social welfare function. Since the lawmaker will choose the welfare-maximizing legal rule in any event, her decision about whether to manipulate individual preferences amounts to a choice between efficient legal rules. The lawmaker prefers legal rules that are easier to enforce and that result in less harm.

II. The Utilitarian Lawmaker

Suppose the lawmaker chooses a utilitarian welfare function. Now she aims to maximize that function, in part by getting the law right. But administering law is difficult. Take the simple, canonical mode of precaution from above. The driver pays damages of D to the pedestrian upon causing harm h. To promote efficient precaution, the driver must fully internalize the pedestrian's harm, meaning D must equal h. If the driver cannot afford D — she is judgment-proof — then she does not fully internalize the harm. She will not take efficient precautions, and the pedestrian will not be fully compensated for an accident. For this reason, liability works better when D is smaller. In addition to the harm to victims that may be compensated through damages, accidents entail costs such as the time and attention of officers, courts, clerks, and jurors.

The lawmaker might reason as follows: to improve liability and reduce costs, we should change the driver's preferences. A *sympathetic* driver who cares about the well-being of pedestrians will take greater precaution and cause fewer and less harmful accidents. In that case, perhaps we can reduce the damages D that are payable to the pedestrian in an accident, mitigating the problem of judgment-proofness, and we can decrease the overall cost of accidents.

---

19 The lawmaker could choose alternatives in between these two, such as a welfare function that includes prosocial preferences but affords them less weight than other preferences. We study the utilitarian social welfare function in Part II because it is an intuitive benchmark and the one most commonly used by economists. We consider the atomistic social welfare function in Part III because there is disagreement about the propriety of including social preferences in the social welfare function.

Does this work? With respect to damages, the answer is no. Liability redistributes costs. Instead of the pedestrian paying, the driver pays. By redistributing the pedestrian's costs to the driver, damages push the driver to take efficient precautions. Sympathy works differently: it *creates* costs, in a sense, by changing what people care about. Without sympathy, the driver's speed causes one harm, which is risk to the pedestrian. With sympathy, the driver's speed causes two harms: risk to the pedestrian, and the psychic harm to the driver of the risk to the pedestrian.[20] The driver internalizes her psychic harm, but she externalizes the pedestrian's actual harm. Because she externalizes harm, the driver — despite her sympathy — takes too little precaution. Sympathy does not generate efficiency.[21] To elicit efficient precaution, the sympathetic driver must be required to pay damages equal to the pedestrian's harm.

Mathematically, adding sympathy "costs" to the driver turns the model from one of unilateral loss to bilateral losses — an accident "harms" both the driver and the pedestrian. Work in this area by Avon Leon, Jennifer Arlen, Robert Cooter, and Ariel Porat explores how the efficiency of various liability rules changes in the case of bilateral losses, including when the losses borne by the injurer are a result of nonlegal sanctions rather than the payment of damages.[22] The important differences between these analyses and ours are

---

20 We can think of prosocial preferences as converting an accident that only harms the victim into an accident that harms the victim and the injurer. *Cf.* Peter A. Diamond, *Single Activity Accidents*, 3 J. Legal Stud. 107 (1974); Jennifer H. Arlen, *Reconsidering Efficient Tort Rules for Personal Injury: The Case of Single Activity Accidents*, 32 Wm. & Mary L. Rev. 41 (1990) [hereinafter Arlen, *Reconsidering*].

21 Scholars studying altruism have made this point. *See, e.g.*, Louis Kaplow, *A Note on Subsidizing Gifts*, 58 J. Pub. Econ. 469 (1995) (explaining that altruists internalize their benefit from giving but not the recipient's benefit from receiving, leading to inefficiency); David D. Friedman, *Does Altruism Produce Efficient Outcomes? Marshall Versus Kaldor*, 17 J. Legal Stud. 1 (1988) (same). *See also* B. Douglas Bernheim & Oded Stark, *Altruism Within the Family Reconsidered: Do Nice Guys Finish Last?*, 78 Am. Econ. Rev. 1034, 1034-35 (1988) ("[I]f A loves B and B is unhappy, then A's love may cause A to be unhappy as well. Worse yet, if B also loves A, then A's love for B would then make B all the more unhappy.").

22 Jennifer H. Arlen, *Re-examining Liability Rules When Injurers as Well as Victims Suffer Losses*, 10 Int'l Rev. L. & Econ. 233 (1990); Arlen, *Reconsidering*, *supra* note 20; Avon Leong, *Liability Rules When Injurers as Well as Victims Suffer Losses*, 9 Int'l Rev. L. & Econ. 105 (1989); Robert Cooter & Ariel Porat, *Does Risk to Oneself Increase the Care Owed to Others? Law and Economics in Conflict*, 29 J. Legal Stud. 19 (2000); Robert Cooter & Ariel Porat, *Should Courts Deduct Nonlegal Sanctions from Damages?*, 30 J. Legal Stud. 401 (2001)

that (1) they consider the effects of imposing losses on an injurer with stable preferences, whereas we consider the effect of changing the injurer's preferences themselves,[23] and (2) we make the injurer's and the victim's utility functions interdependent by making the injurer care about the victim's well-being.

A simple example proves the point. The pedestrian's expected costs equal $p(x)(h - D)$, and the sympathetic driver's expected costs equal $x + p(x)(D + \alpha h - \alpha D)$, where $\alpha > 0$ indicates the sympathy weight that the driver attaches to the pedestrian's utility. Note that the driver fully sympathizes with the pedestrian, caring about both the harm and the damages from the pedestrian's perspective. Social costs are given by $x + p(x)(h + \alpha h - \alpha D)$. To elicit efficient precautions, the driver's costs must align with society's costs. This requires setting D equal to h — just as in the case of the selfish driver. No matter how much the driver sympathizes with the pedestrian, D must equal h, otherwise the driver will take too little precaution.[24]

We have shown that sympathy (even strong sympathy) does not reduce the damages needed to induce efficient precaution. But perhaps sympathy is the wrong medicine. As we conceive of it, sympathy incorporates the overall well-being of the pedestrian into the driver's utility function. When the pedestrian bears harm, the driver pays a psychic cost, but when the pedestrian receives compensation, the driver enjoys a psychic benefit. For precautions, then, sympathy has cross-cutting effects. The pedestrian's harm encourages the driver to take more precaution, but the damages payable to the pedestrian encourage her to take less precaution.

Instead of sympathy, suppose the lawmaker inculcates *harm aversion*. This means that the driver has a distaste for imposing harm but does not fully

---

[hereinafter Cooter & Porat, *Should courts deduct?*]. This literature examines the effects of bilateral precaution (in addition to bilateral loss) and risk spreading from insurance, as well as negligence liability regimes. We do not address these interesting extensions here, but we believe that they are a fruitful direction for future work. We focus on the particular effects of changing preferences rather than just adding costs to the canonical model.

23 For example, Cooter and Porat write that nonlegal sanctions that do not confer benefits on third parties "destroy value." Cooter & Porat, *Should courts deduct?*, *supra* note 22, at 414. This cannot be said of a driver who develops sympathy for her victim and who thereby changes her preferences, because value is defined by reference to preferences.

24 The general recursive problem created by other-regarding preferences remains if the driver's preferences are to maximize social welfare. Suppose that the driver's utility function is f and the pedestrian's utility function is g. Social welfare is given by $w = f + g$. If the driver seeks to maximize social welfare so that $f = f + g$ then $w = 2f + g$, but then the driver is no longer maximizing social welfare.

identify with the well-being of the pedestrian (thereby ignoring the benefit received by the pedestrian from compensation). In our model, the harm-averse driver's cost function would be given by $x + p(x)(D + \alpha h)$. As above, the pedestrian's costs equal $p(x)(h - D)$, so social costs would be given by $x + p(x)(h + \alpha h)$. To elicit efficient precautions, the driver's costs must align with society's costs. Yet again, this requires setting D equal to h.

In sum, it does not matter whether the driver is selfish, sympathetic, or harm averse. In all three cases, the lawmaker must set damages equal to the harm of an accident to induce efficient precaution. With sympathy or harm aversion, the driver's activity causes two harms: risk to the pedestrian, and psychic harm to the driver. The driver internalizes her psychic harm, but she externalizes the pedestrian's actual harm. To correct the driver's incentives, liability must force the driver to internalize the pedestrian's actual harm. That harm is the same whatever the driver's preferences.

This result has another implication. We have not thus far discussed how the lawmaker might use the law to change people's preferences. The amount of damages seems like a natural possibility. Through damages, the lawmaker can convey social disapproval of harm creation and possibly cause people to become more sympathetic or harm averse. For this reason, setting damages at a punitively high level might seem like a good idea. What the foregoing result shows, however, is that setting D at any level other than h will be inefficient *no matter how effectively it performs its expressive function*.

Recall that the lawmaker has two objectives for any efficient legal rule: mitigating judgment-proofness to enhance the enforceability of that rule, and lowering the overall costs of accidents. We have shown that preference change does not affect optimal damages, so it cannot do much for judgment-proofness. What about overall accident costs? If $D = h$, then harm-averse people will drive more carefully than selfish people.[25] This is intuitive: selfish people are trying to avoid damages, while harm-averse people are trying to avoid damages *and* harm. Consequently, inculcating harm aversion will lower the number and severity of accidents. However, this reduction does not come without complications. To reduce accidents, the lawmaker must create psychic costs from the imposition of harm, and these psychic costs must be evaluated.

We have analyzed the precaution of selfish and harm-averse drivers. What about sympathetic drivers? Interestingly, sympathetic drivers are no more

---

25 If D equals h, drivers will take socially efficient precaution. The socially efficient precaution in the case of a harm-averse driver satisfies $p'(x^*) = \frac{1}{h(1+\alpha)}$, whereas socially efficient precaution in the case of a selfish driver satisfies $p'(x^*) = \frac{1}{h}$. Note that different drivers take different amounts of precaution, but the amount is efficient in all cases.

careful than their selfish counterparts.[26] The sympathetic driver does not suffer any net psychic disutility from the harm she imposes on the pedestrian because the compensatory damages she pays make the pedestrian, and therefore the driver herself, whole. For the lawmaker, sympathy has no benefit; it neither reduces optimal damages nor encourages more precaution.

III. The Atomistic Lawmaker

So far we have assumed that the lawmaker is utilitarian, giving equal weight to all individuals' preferences, including preferences one has for the well-being of others. Now imagine instead a different, hardheaded lawmaker who seeks to inculcate sympathy in drivers. When her adviser notes that sympathy introduces a new psychic cost, complicating the social welfare calculation, the lawmaker responds that she does not care one wit about the driver's feelings, but only about the pedestrian's safety. In other words, impose a psychic harm on the driver but ignore that harm in social welfare. The lawmaker adopts an *atomistic* welfare function. She views prosocial preferences instrumentally and disregards any social welfare that flows from those preferences.

It is tempting to reject atomistic welfare functions. This is because using one requires making difficult philosophical judgments. Choosing which preferences to exclude is equivalent to sorting "better" and "worse" preferences. As we have already explained, economics has no tool for making that choice.[27] Not everyone is troubled by this challenge, however. Scholars[28] (including some economists[29]) have argued that certain benefits should be excluded

---

26 The sympathetic driver's cost function equals $x + p(x)(D + \alpha h - \alpha D)$. Assuming D equals h , then she takes socially efficient precaution, which satisfies $p'(x^*) = \frac{1}{h}$.

27 Recall that some economists have tried to develop such a tool. *See* Weisbrod, *supra* note 16.

28 In law, see, for example, Joseph William Singer, *Normative Methods for Lawyers*, 56 UCLA L. Rev. 899, 919-20 (2009) ("[S]ome preferences are simply intolerable and do not enter into the calculus of conscientious judges and legislators."); Daphna Lewinsohn-Zamir, *The Objectivity of Well-Being and the Objectives of Property Law*, 78 N.Y.U. L. Rev. 1669, 1682 (2003) ("Some writers suggest that certain objectionable preferences, such as those based on cruelty, racism, and prejudice, should be disregarded in social policy.").

29 *See, e.g.*, John C. Harsanyi, *Rule Utilitarianism and Decision Theory*, 11 Erkenntnis 25, 30 (1977) (arguing that scholars should "define social utility in terms of the various individuals' 'true' preferences," thus disregarding "not only preferences distorted by factual or logical errors, but also preferences based on clearly antisocial attitudes, such as sadism, resentment, or malice"); George J. Stigler,

from welfare. They focus on the "benefits" derived from acts like murder, rape, racism, and cruelty. But this approach is not necessarily sufficient. In addition to excluding antisocial benefits, there may be advantages to excluding prosocial feelings as well.[30]

To see why, let's return to the model. The pedestrian's costs equal $p(x)(h - D)$. If the driver is selfish, her costs equal $x + p(x)D$, and social costs are the sum of the two: $x + p(x)h$. In this scenario, no one has prosocial preferences, so the welfare function is the same for both the atomistic and utilitarian lawmaker. To minimize social costs, damages D must equal harm h.

But what if the lawmaker inculcates sympathy in the driver? The pedestrian's costs do not change, but the driver's costs become $x + p(x)(D + \alpha h - \alpha D)$, where $\alpha > 0$ again indicates the sympathy weight. The atomistic lawmaker ignores the driver's concern for the well-being of the pedestrian, meaning she excludes sympathy from the welfare function. Consequently, social costs are given by $x + p(x)h$. The driver will choose the efficient level of precaution if $D = h$, meaning compensation is perfect. But suppose compensation is imperfect. In fact, suppose there is no compensation at all. The driver will *still* choose the efficient level of precaution if $\alpha = 1$, meaning she is perfectly sympathetic.

In this scenario, sympathy eliminates the need for damages. A perfectly sympathetic driver will behave efficiently whether she is liable for all, some, or none of the harm she causes. Sympathy solves the problem of judgment-proofness and the costs of assessing and collecting damages more generally.

This happy result has three limitations. First, it only achieves one of the lawmaker's objectives. Recall that the lawmaker wants to reduce both damages and the overall harm from accidents. Perfect sympathy eliminates the need for damages, but it does not cause the driver to take more precaution. The perfectly sympathetic driver internalizes a (psychic) cost that equals the (damages) cost internalized by the selfish driver. They both internalize the same amount in the event of an accident, so they both take the same precaution

---

The Optimum Enforcement of Laws, 78 J. Pol. Econ. 526, 527 (1974) ("what evidence is there that society sets a positive value upon the utility derived from a murder, rape, or arson?").

30 Other scholars have proposed excluding beneficent other-regarding preferences from welfare, though not for the reason we discuss. *See, e.g.*, John C. Harsanyi, *Problems with Act-Utilitarianism and with Malevolent Preferences*, *in* Hare and Critics: Essays on Moral Thinking 89, 97-98 (Douglas Seanor & N. Fotion eds., 1988) ("[I]t seems to me that even socially *desirable* external preferences should be *excluded* from our social-utility function.").

to avoid it.[31] Since they both take the same level of precaution, the total harm from accidents remains the same.

Second, sympathy must be perfect. If the driver's sympathy weight is less than 1, even just barely, achieving efficiency requires perfect compensation.[32] To achieve efficiency, the sympathetic driver must internalize the harm she causes. In reality, the driver internalizes the damages she pays and (here is the psychic part) the pedestrian's *uncompensated* harm. For a perfectly sympathetic driver, the compensation she pays and the uncompensated harm equal the actual harm. But for a less-than-perfectly sympathetic driver this does not hold. The driver does not place enough weight on the uncompensated harm. To offset that, the compensation the driver pays must rise, but as soon as it does the uncompensated harm matters even less. For the less-than-perfectly sympathetic driver, the compensation she pays and the uncompensated harm equal the actual harm *only* when compensation is perfect. (Likewise, if the driver's sympathy weight exceeds 1, achieving efficiency again requires perfect compensation.)

The result has a third limitation: the lawmaker must adopt an atomistic social welfare function. She must impose a cost in the form of prosocial guilt on the driver individually, but she must ignore that cost collectively.[33] This is not a quirk of our model. It is a general observation about the logic of — and a fallacy in — common arguments about preference change. Eliciting prosocial preferences can reduce some social costs, where "social costs" are assessed in terms of the new, post-preference-change welfare function. But achieving that outcome requires excluding from welfare those prosocial preferences. This is the virtue paradox. For virtue to improve welfare, we must ignore its toll on the virtuous.

---

31 The sympathetic driver's cost function equals $x + p(x)(D + \alpha h - \alpha D)$. Assuming $D = 0$ and $\alpha = 1$, the driver takes precaution that satisfies $p'(x^*) = \frac{1}{h}$. The selfish driver's cost function equals $x + p(x)(D)$. Assuming $D = h$, the driver takes precaution that satisfies $p'(x^*) = \frac{1}{h}$.

32 Efficiency is achieved when the harm the sympathetic driver causes equals the disutility the sympathetic driver bears: $h = D + \alpha h - \alpha D$. If $\alpha = 1$, the equation holds regardless of the values of D and h. If $\alpha \neq 1$, the equation holds only if $h = D$.

33 That the cost is psychic does not, by itself, justify ignoring it. *See* Kaplow & Shavell, *supra* note 4, at 1343 n.931 ("Although the distinction between tangible and psychic externalities may make sense as a matter of policy, . . . it is difficult to understand the distinction as a matter of first principle. After all, everything that affects someone's well-being is ultimately perceived by the senses and mediated in the human brain; it is unclear what is the a priori normative basis for expressing concerns about certain triggers of particular neurons over others.").

Suppose the lawmaker dislikes these limitations. Can she improve matters by inculcating harm aversion instead of sympathy? The harm-averse driver's costs equal $x + p(x)(D + \alpha h)$, and the pedestrian's costs equal $p(x)(h - D)$ as usual. The atomistic lawmaker ignores prosocial preferences, so social costs are given by $x + p(x)h$. To achieve efficiency, the driver's costs must match social costs. This requires the driver to pay damages of $D = h(1 - \alpha)$. Optimal damages decrease as α increases. Thus, harm-aversion mitigates the problem of judgment-proofness.

This result is more intuitive than it might seem. To induce the driver to take the socially optimal amount of precaution, the driver must internalize the harm she causes to the pedestrian. Since the driver already internalizes a psychic, weighted amount of the harm she causes to the pedestrian, the efficient amount of damages only needs to make up the difference between this psychic quantity and the actual harm to the pedestrian. As the psychic weight the driver attaches to the harm increases, compensation must fall for the equation to remain in balance. More harm aversion implies lower optimal damages.[34]

Unlike sympathy, harm aversion need not be perfect — α need not equal 1 — to reduce optimal damages. This makes inculcating harm aversion a more robust strategy. If the lawmaker can make the driver even a little harm-averse, optimal damages decrease. However, harm aversion is not a magic bullet. The harm-averse driver takes the same precaution as the sympathetic and selfish driver.[35] Consequently, the total number of accidents and the associated costs remain the same. Furthermore, harm aversion implicates the virtue paradox. The lawmaker must ignore in social welfare the individual, psychic harm that she worked to inculcate.

IV. Preferences or Commitments?

For a lawmaker intent on changing individuals' preferences, our analysis is clear: inculcate harm aversion. Given a utilitarian social welfare function, harm aversion generates more precaution than sympathy or selfishness, even though it has no effect on the optimal amount of damages. Given an atomistic

34 This only works for an atomistic welfare function. Given a utilitarian welfare function, efficiency would require the driver to internalize the harm she causes to the pedestrian *and* the psychic harm she causes to herself. To internalize all of that requires setting D equal to h.

35 The harm-averse driver's cost function equals $x + p(x)(D + \alpha h)$. Damages are optimal when $D = h(1 - \alpha)$. Assuming optimal damages, the harm-averse driver takes precaution that satisfies $p'(x^*) = \frac{1}{h}$.

social welfare function, harm aversion reduces optimal damages,[36] even though it has no effect on how much precaution is taken. Whatever welfare function one selects, harm aversion outperforms the alternatives. However, harm aversion — like sympathy — comes with a downside: psychic cost. The utilitarian lawmaker accepts that cost in exchange for more precaution. The atomistic lawmaker simply ignores that cost. In both cases preferences are changing, so economics cannot guide any comparisons. The lawmaker must make philosophical choices, not economic choices.

Our conclusion might raise optimism about harm aversion but discouragement about law's expressive function in general. Yes, inculcating harm aversion helps, but only if we accept strong and contestable moves, like embracing the virtue paradox. But perhaps we need not accept these moves. Perhaps we can profit from law's expressive function by reconsidering the nature of preferences.

What if there is space between choices and preferences? Sometimes people might make choices not because doing so helps them best realize their goals or because of sympathy for others, but because of habit or moral commitment. To demonstrate, people might conform to the principle "treat all people with equal respect" not because it helps them achieve a particular outcome, but because the principle delimits moral conduct or structures an identity. This idea tracks Amartya Sen's distinction between sympathy and commitment.[37] This distinction is controversial,[38] and we do not defend it here. Instead, we develop a two-step conjecture. First, one understanding of law's value-shaping function is that it forms patterns of action or modes of reasoning that operate beyond individual preference satisfaction. And second, this might make preference change a more attractive strategy.

Suppose the lawmaker is utilitarian. As discussed above, efficient damages would ordinarily be the same for selfish, sympathetic, and harm-averse drivers alike: $D = h$. Suppose, however, that the lawmaker can cause drivers to act *as if* they were sympathetic, even if they are not. In that case, all drivers choose precaution to minimize costs given by $x + p(x)(D + \alpha h - \alpha D)$. If a driver is

---

36 Recall that perfect sympathy also reduces (in fact eliminates) damages, but this seems like a rare case. Less-than-perfect sympathy has no effect on optimal damages.

37 *See* Amartya Sen, *Why Exactly is Commitment Important for Rationality?*, 21 Econ. & Phil'y 5 (2005); Amartya Sen, *Goals, Commitment, and Identity*, 1 J.L. Econ. & Org. 341 (1985); Amartya K. Sen, *Rational Fools: A Critique of the Behavioral Foundations of Economic Theory*, 6 Phil'y & Pub. Affairs 317 (1977).

38 Elias L. Khalil, *Sentimental Fools: A Critique of Amartya Sen's Notion of Commitment*, 40 J. Econ. Behavior & Org. 373 (1999).

actually selfish, then the social costs of her driving are given by $x + p(x)h$, and efficient damages remain $D = h$. If a driver is actually sympathetic, then social costs are given by $x + p(x)(h + \alpha h - \alpha D)$, and efficient damages also remain $D = h$. However, if the driver is actually harm averse, then social costs are given by $x + p(x)(h + \alpha h)$. Efficiency requires the driver's cost to equal society's costs, which requires setting $D = \frac{h}{1-\alpha}$. Since $\alpha > 0$, encouraging harm-averse drivers to behave as if they were sympathetic *increases* the amount of damages necessary to induce efficient precaution.[39] Moreover, whether the driver is actually selfish, sympathetic, or harm averse, causing her to behave as if she were sympathetic has no effect on the amount of precaution that she takes, assuming the lawmaker assigns the efficient amount of damages.[40] Encouraging drivers to act as if they were sympathetic is not promising.

Alternatively, suppose the lawmaker can cause drivers to act as if they are harm averse. All drivers will choose x to minimize $x + p(x)(D + \alpha h)$. If a driver is actually selfish, then social costs are given by $x + p(x)h$, and the efficient damages rule becomes $D = h(1 - \alpha)$. This is a good and surprising outcome. Although making a selfish driver harm averse does not reduce optimal damages,[41] making a selfish driver behave *as if* she is harm averse does. The key is that damages and the psychic harm from an accident act as substitutes in this case. Imagine a selfish driver facing efficient damages of $D = h$. As she begins to behave as if she is harm averse, she takes greater precaution than she otherwise would *and greater precaution than is socially optimal*, because her harm aversion is not included in social welfare (recall that the driver is actually selfish). The lawmaker can benefit from this increase in precaution by reducing the amount of damages it imposes.

---

39 Here is the intuition: making the harm-averse driver behave as if she were sympathetic causes her to act as if she derives some benefit from the damages that she pays to the pedestrian in the event of an accident (this is the difference between a sympathetic driver and a harm-averse driver). This benefit reduces the (psychic) cost to the driver of accidents, encouraging her to take fewer precautions. Damages must increase to reestablish the efficient amount of precaution.

40 Selfish drivers who act selfishly and sympathetic drivers who act sympathetically take precaution to satisfy $p'(x^*) = \frac{1}{h}$. *See supra* notes 25, 26. For selfish and sympathetic drivers who act sympathetically, efficient damages equal h and, assuming efficient damages, both drivers again take precaution to satisfy $p'(x^*) = \frac{1}{h}$. Harm-averse drivers take precaution to satisfy $p'(x^*) = \frac{1}{h(1+\alpha)}$. *See supra* note 25. For harm-averse drivers who act as if they are harm averse, efficient damages equal $\frac{h}{1-\alpha}$, and assuming efficient damages, the harm-averse driver again takes precaution to satisfy $p'(x^*) = \frac{1}{h(1+\alpha)}$.

41 *See supra* Part III.

What if drivers are actually sympathetic? In that case, social costs are given by $x + p(x)(h + \alpha h - \alpha D)$ and efficient damages are set at $D = \frac{h}{1+\alpha}$, which is of course less than h. This is also a good outcome. Causing a sympathetic driver to act as if she is harm averse reduces optimal damages. This is because the deterrent effect of damages on the driver is no longer diminished by the benefit the driver gets from compensating the pedestrian. Increasing the deterrent effect of damages allows the lawmaker to reduce the amount of damages that must be imposed to induce efficient precaution.

What if the driver is actually harm averse? She will choose x to minimize $x + p(x)(D + \alpha h)$. Social costs are given by $x + p(x)(h + \alpha h)$, and the efficient damages rule remains $D = h$.

What effect, if any, does all of this have on the amount of precaution taken by these drivers? For the selfish driver, acting as if she is harm averse and facing efficient damages yields the same precaution as if she simply acted selfishly and faced efficient damages.[42] For the harm-averse driver, acting harm averse does not change the precaution she takes, so the analysis matches that above.[43] However, the sympathetic driver who acts like a harm-averse driver and faces efficient damages chooses precaution x to minimize social costs of $x + p(x)(h + \alpha h - \frac{\alpha h}{1+\alpha})$. Recall that she would choose precaution to minimize social costs of $x + p(x)h$ if she instead faced damages of $D = h$, which would be efficient if she acted sympathetically instead of harm averse. As long as $\alpha > 0$, then she will choose more precaution if she is induced to act harm averse than if not.

In sum, inculcating a rule to behave as if one were harm averse will reduce optimal damages for selfish and sympathetic drivers, while having no effect on optimal damages for actual harm-averse drivers. This strategy is especially attractive in the case of selfish drivers. We can reduce the optimal damages they pay, and mitigate the problem of judgment-proofness, *without* requiring the driver to suffer psychic harm or the lawmaker to embrace the virtue paradox. Furthermore, this same strategy reduces the total harm from accidents caused by sympathetic drivers. It is the only approach that promises

---

42 The selfish driver acts as if her cost function equals $x + p(x)(D + \alpha h)$. Damages are optimal when $D = h(1 - \alpha)$. Assuming optimal damages, the selfish driver takes precaution that satisfies $p'(x^*) = \frac{1}{h}$. This is the same precaution she would take if she acted selfishly. *See supra* note 25.

43 The harm-averse driver acts as if her cost function equals (and in fact it does equal) $x + p(x)(D + \alpha h)$. Damages are optimal when $D = h$. Assuming optimal damages, the harm-averse driver takes precaution that satisfies $p'(x^*) = \frac{1}{h(1+\alpha)}$. This is the same precaution the harm-averse driver would take in the original analysis. *See supra* note 25.

to both reduce the damages needed to induce efficient precaution and reduce the amount of harm from driving.

As is obvious by this point, the difference between sympathy and harm aversion in this context has only to do with the different way that these preferences treat the receipt of damages by the pedestrian. The disutility from an accident borne by a sympathetic driver is $D + \alpha h - \alpha D$, where the disutility borne by the harm-averse driver is only $D + \alpha h$. Comparing these costs reveals that a sympathetic driver will have the same incentives to take precaution as the harm-averse driver if, upon the occurrence of an accident, the driver is liable for a fine payable to the government or some other sanction that does not benefit the pedestrian. This means that, whereas inculcating an aversion to harm requires an intervention into the preferences or commitments of individuals, a sympathetic driver can be made functionally harm averse simply by changing the nature of the sanction.

In the case of *homo economicus*, the only relevant feature of a sanction is the amount of harm it imposes on the actor herself. As soon as we allow for prosocial preferences, it opens up alternative points of leverage over the actor's conduct. One of us has written about the particular benefits of damages as a deterrent for individuals who are motivated by animus (i.e., they derive utility from the objects of their animus being made worse off).[44] In that context, damages have a multiplicative deterrent effect: paying damages directly makes the actor worse off by depriving her of the consumption that she would otherwise be able to enjoy, and paying them to the object of her animus makes her yet worse off because she dislikes helping those she wants to see harmed. In the context of prosocial preferences, we see the opposite. From an efficient pricing perspective, it is better to make the receipt of the damages by the pedestrian less salient or replace damages with fines or hard treatment.

The rule for optimal damages in this case is also the same as in the case of a utilitarian lawmaker who induces a selfish driver to act as if she were harm averse. This is because a utilitarian lawmaker regulating selfish drivers has the same social cost function as an atomistic lawmaker (regardless of driver preferences). The difference is that whereas the atomistic lawmaker can manipulate either the driver's actual preferences or moral commitments to induce her to act like a harm-averse driver, the utilitarian lawmaker cannot change the driver's preferences without changing the form of her social cost function, and so she must rely on the habit-forming or commitment-creating

---

44 *See* Andrew T. Hayashi, *The Law and Economics of Animus*, U. Chi. L. Rev. (Forthcoming 2022).

effects of the law to substitute for the amount of damages needed to induce efficient precaution.

Conclusion

The notion that good laws can make for a good society, not through a system of carrots and sticks but by expressing good values that improve individual character, has wide currency. On this account, we may be able to eliminate, or at least reduce, the various forms of expensive coercion needed to steer the "bad man" to socially desirable conduct. It is an optimistic account. We have highlighted the difficulties that the expressive function runs into within a welfarist framework. Not only does a project of changing individual preferences run into the familiar problem of making interpersonal utility comparisons, but it may not change optimal damages or precaution.

We have identified two ways to realize the benefits of law's expressive function. The first is to use the law to inculcate norms or moral commitments that are not reflective of individual preferences. The second is to embrace an atomistic welfare function that ignores prosocial preferences. Both solutions require loosening some of the traditional assumptions of positive or normative economics, but so too does the mere possibility of changing values or preferences through the expressive function.

Appendix

This appendix summarizes the results reported in the body of the Article. We consider an individual i engaged in an activity that imposes monetary harm h on a third party j with probability $p(x)$, where x is the amount of precaution taken by i. The probability function has properties $p'(x) < 0 < p''(x)$. We assume that the marginal cost of precaution is 1, and that if the harm occurs then i is strictly liable for monetary sanction D, which is transferred to j. The socially optimal amount of precaution x^* is chosen if the costs to i from the creation of harm are equal to the social costs of the harm. We assume that the cost of imposing a sanction D is increasing in the amount of the sanction.[45]

We consider first the case of a **utilitarian social cost function**, in which social costs are equal to the sum of the costs borne by i and j from the activity.

- **Selfish.** In the baseline case, i is selfish, so she chooses x to minimize $x + p(x)D$. Social costs are therefore $x + p(x)h$. Private costs are equal to social costs only if $D = h$.
- **Sympathy.** Suppose that i cares about the utility of j, and assigns weight α to j's well-being. In that case, i chooses x to minimize $x + p(x)(D + \alpha h - \alpha D)$. Social costs are equal to $x + p(x)(h + \alpha h - \alpha D)$. Again, private costs are equal to social costs only if $D = h$.
- **Harm aversion.** Suppose instead that i cares about the harm that she imposes on j, but either does not value the benefit that j gets from D, or the monetary sanction is a fine that is paid to the government (not to j). In that case, i chooses x to minimize $x + p(x)(D + \alpha h)$. Social costs are equal to $x + p(x)(h + \alpha h)$. Again, private costs are equal to social costs only if $D = h$.

In all three cases, the amount of damages required to induce efficient precaution is the same and equal to the harm caused. Next, we consider what happens if i can be induced to act *as if* she had preferences other than those that she has, perhaps out of habit or a moral commitment.

- If i has selfish preferences, then social costs are $x + p(x)h$.
  - If i can be induced to act as if she is sympathetic, then she chooses x to minimize $x + p(x)(D + \alpha h - \alpha D)$. Private costs are equal to social costs only if $D = h$.

---

45 This could be because of judgement-proofness concerns. D would represent the expected damages, which depend on the probability of enforcement. It is costly to increase the probability of enforcement.

  - If i can be induced to act as if she is harm averse, then she chooses x to minimize $x + p(x)(D + \alpha h)$. Private costs are equal to social costs only if $D = h(1 - \alpha)$.

- If i is sympathetic, then social costs are $x + p(x)(h + \alpha h - \alpha D)$.
  - If i can be induced to act as if she is harm averse, then she chooses x to minimize $x + p(x)(D + \alpha h)$. Private costs are equal to social costs only if $D = \frac{h}{1+\alpha}$.
  - If i can be induced to act as if she is selfish, then she chooses x to minimize $x + p(x)D$. Private costs are equal to social costs only if $D = h$.
- If i is harm averse, then social costs are $x + p(x)(h + \alpha h)$.
  - If i can be induced to act as if she is sympathetic, then she chooses x to minimize $x + p(x)(D + \alpha h - \alpha D)$. Private costs are equal to social costs only if $D = \frac{h}{1-\alpha}$.
  - If i can be induced to act as if she is selfish, then she chooses x to minimize $x + p(x)D$. Private costs are equal to social costs only if $D = h(1 + \alpha)$.

We consider next the case of an **atomistic social cost function** that ignores other-regarding preferences, which implies that the social cost function is equal to $x + p(x)h$.

- If i is sympathetic, she will choose x to minimize $x + p(x)(D + \alpha h - \alpha D)$. Social costs are equal to private costs only if $D = h$.
- If i is harm averse (or the monetary sanction is a fine), she will choose x to minimize $x + p(x)(D + \alpha h)$. Social costs are equal to private costs only if $D = h(1 - \alpha)$.