

University of Chicago Law School

Chicago Unbound

Journal Articles

Faculty Scholarship

1998

Strategic Basis of Principled Behavior: A Critique of the Incommensurability Thesis

Eric A. Posner

Follow this and additional works at: https://chicagounbound.uchicago.edu/journal_articles



Part of the [Law Commons](#)

Recommended Citation

Eric Posner, "Strategic Basis of Principled Behavior: A Critique of the Incommensurability Thesis," 146 University of Pennsylvania Law Review 1185 (1998).

This Article is brought to you for free and open access by the Faculty Scholarship at Chicago Unbound. It has been accepted for inclusion in Journal Articles by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.

THE STRATEGIC BASIS OF PRINCIPLED BEHAVIOR:
A CRITIQUE OF THE INCOMMENSURABILITY THESIS

ERIC A. POSNER[†]

When my love swears that she is made of truth
I do believe her though I know she lies,
That she might think me some untutored youth
Unlearnèd in the world's false subtleties.
Thus vainly thinking that she thinks me young,
Although she knows my days are past the best,
Simply I credit her false-speaking tongue;
On both sides thus is simple truth suppressed.
But wherefore says she not she is unjust,
And wherefore say not I that I am old?
O, love's best habit is in seeming trust,
And age in love loves not to have years told.
Therefore I lie with her, and she with me,
And in our faults by lies we flattered be.

—Shakespeare's *Sonnet 138*

INTRODUCTION

The incommensurability thesis holds that people cannot always value options along a common metric that is normatively justified. Most advocates of this thesis argue that people can choose among options, but that the choice depends on qualitative differences between options that cannot be reduced to vectors on a single dimension of evaluation.¹ The thesis, it should be noted, is not a theory of choice.

[†] Professor of Law, University of Pennsylvania. Thanks to Matthew Adler, Richard Craswell, Peter Huang, Lewis Kornhauser, David Strauss, Cass Sunstein, and Symposium participants for very helpful comments.

¹ For modern discussions, see, for example, ELIZABETH ANDERSON, *VALUE IN ETHICS AND ECONOMICS* (1993); JOSEPH RAZ, *THE MORALITY OF FREEDOM* 321-66 (1986); Matthew Adler, *Incommensurability and Cost-Benefit Analysis*, 146 U. PA. L. REV. 1371 (1998); Ruth Chang, *Introduction*, in *INCOMMENSURABILITY, INCOMPARABILITY, AND PRACTICAL REASON I* (Ruth Chang ed., 1997); Richard A. Epstein, *Are Values Incommensurable, or Is Utility the Ruler of the World?*, 1995 UTAH L. REV. 683; Martha C. Nussbaum, *Flawed Foundations: The Philosophical Critique of (a Particular Type of) Economics*, 64 U. CHI. L. REV. 1197, 1199-203 (1998); and Cass R. Sunstein, *Incommensurability and Valuation in Law*, 92 MICH. L. REV. 779 (1994). An interesting early discussion, unfortunately ignored in the modern literature but full of relevant insights and anticipations,

Rather, it is a label attached to a social phenomenon: the tendency of many people to refuse to make tradeoffs in everyday life that are said to be demanded by the theory of rational choice. Because these refusals are anomalies for rational choice theory, they are said to justify rejection of that theory and its replacement with a superior theory of choice. Conceived in this way, the incommensurability thesis justifies a research program devoted to the discovery of a theory that unifies our intuitions and observations about the choices people make.

It is a difficult question when and whether to abandon or modify an old theory because it does not fully explain the phenomena it is intended to explain. Usually, people do not abandon an old theory until a better theory has presented itself. Although some advocates of the incommensurability thesis have presented their own theories of choice,² these theories have little explanatory power. Right now, then, the question is whether incommensurability poses a serious enough challenge to rational choice to justify either abandoning it or modifying it. I want to make only a modest claim: One should not reject rational choice theory because of its failure to explain incommensurability without first exploring how rational choice theory might deal with incommensurability. That is to say, does the incommensurability thesis really identify anomalies in rational choice theory?

I will argue that it does not.³ Although the incommensurability thesis often describes people's *representations* about themselves, it does not describe their actual *behavior*—that is, the choices they make in everyday life—although their representations sometimes influence their behavior. People rationally make incommensurability claims in order to obtain strategic advantages in their interactions with others. Incommensurability claims do not reflect people's interests and values; they conceal them.

The argument has two parts. First, I present a model in which people take actions and make statements in the hope of acquiring a valuable reputation for trustworthiness. In this model, false state-

can be found in GEORG SIMMEL, *THE PHILOSOPHY OF MONEY* (Tom Bottomore & David Frisby trans., 1978).

² See, e.g., ANDERSON, *supra* note 1, at 17 (describing an "expressive theory" of rational action).

³ The same question is posed, but the opposite answer is given, in Richard Warner, *Impossible Comparisons and Rational Choice Theory*, 68 S. CAL. L. REV. 1705 (1995) (arguing that incommensurability lies behind behaviors, such as voting and cooperation, that are said to be inexplicable under rational choice theory).

ments about one's values emerge in equilibrium, and these false statements have a common element: They take the form of principles. Whether or not people believe that they are principled, most people will violate their principles when doing so serves their interests, even though they must and do publicly express their commitment to those principles. Second, I argue that incommensurability claims are a type of principled claim. Incommensurability claims emerge in equilibrium even though people do not act consistently with them and even though they do not express people's values. Incommensurability is not so much a problem about values; it is a problem about social coordination and ideology. The Article concludes with a discussion of the normative implications of the analysis, arguing that incommensurability claims should have little weight in any theory of social choice. As a positive matter, however, governments must and do give weight to incommensurability claims.

To motivate intuitions, I set out some examples of incommensurability, all taken from Sunstein.⁴

- A person cancels lunch with a friend and feels bad, but he could not properly offer the friend \$20⁵ as compensation. Friendship and money are not commensurable.
- An employer could offer an employee an extra \$1000 to do work that would require an absence from home for a month, but could not, without insulting the employee, offer to pay the employee \$1000 to persuade him to work away from home for a month. Family relations and money are not commensurable.⁶
- A beautiful mountain may strike awe in the heart of an observer, but a statement like "this mountain is worth \$10 million" would not escape from his lips. Environmental wonders and money are not commensurable.⁷
- One should not offer to pay a neighbor \$20 to mow one's lawn, or offer to pay a person for sex. Neighborly and intimate relationships are not commensurable with money.⁸

⁴ See Sunstein, *supra* note 1.

⁵ See *id.* at 785.

⁶ See *id.* at 786; see also RAZ, *supra* note 1, at 348-49.

⁷ See Sunstein, *supra* note 1, at 786-87.

⁸ See *id.* at 787.

- It is insulting to suggest that an actor's or teacher's worth is equal to his pay or any amount of monetary compensation.⁹
- It is wrong to say that a worker's life is worth eight million dollars (or some other amount), even though the worker accepts a wage premium of \$800 for a job that carries with it a one-in-ten-thousand chance of fatal injury above the norm.¹⁰

These examples present the following puzzle: Although the statements command a great deal of assent and appeal to one's intuitions, and the simplest explanation for their force is that certain options are incommensurable, people in fact behave as though they believed that these options are commensurable.

The reason that people treat supposedly incommensurable options as though they were commensurable is that resources are scarce and choices must be made. Sunstein mentions the "desperately poor [person]" who might abandon a friend for a sum of money, but he is altogether too delicate.¹¹ Lawyers, businessmen, movie moguls, politicians, and even ordinary people constrained by time and resources drop friends when the costs of maintaining the friendship become too high. People routinely risk or violate family relationships by accepting attractive job offers in distant locations. Individuals make implicit valuations of the environment when they choose detergents, buy paper products, recycle newspapers but not bottles, purchase large houses rather than small apartments, and litter. People pay for sex or for companionship, sometimes overtly and other times in carefully disguised manners, such as in the form of gifts. Artists and teachers become investment bankers when their pay falls below a certain level. Workers accept premia for risks, and agencies use these risk premia in order to calculate the costs and benefits of regulations.

The puzzle, then, is why people say, and even believe, that certain options are incommensurable and sometimes act in ways designed to maintain the pretense of their incommensurability, while at the same time acting as though those options were commensurable. Sunstein, Anderson, and others seize one horn of the dilemma by arguing that the incommensurability of options is fundamental, and that a theory of choice that accounts for it can explain people's actions.¹² I will

⁹ See *id.* at 788.

¹⁰ See *id.* at 804.

¹¹ *Id.* at 813.

¹² See *supra* note 1.

seize the other horn of the dilemma and argue that "incommensurability claims," as I call them, emerge in an equilibrium in which people rationally seek partners for the purpose of obtaining gains from cooperation.

I. A MODEL OF PRINCIPLED BEHAVIOR

People can obtain strategic advantages by limiting their options prior to interactions with others. Suppose, for example, that a buyer and a seller are negotiating over the sale of a widget, and the price will be between five and ten. Both parties are willing to accept any price in that range, but the buyer wants the lowest price and the seller wants the highest price. If the buyer can bind himself against accepting a high price, and the seller knows this, the buyer has a strategic advantage. One way the buyer might be able to bind himself to a low price is by making a deal with a third party prior to this negotiation, in which the buyer promises to pay the third party twenty if he pays more than five for the widget. Confronted by this deal, the seller would concede to the price of five, knowing that the buyer would not accept a price of six or higher, since that would require the buyer to pay twenty-six or more for something that he values at ten. This kind of strategic maneuver can be called a *commitment mechanism*.¹⁵

One uses commitment mechanisms not only for obtaining advantages against competitors, but also for giving assurances to friends or partners. Suppose that the seller and the buyer agree on a price, but although the seller cannot deliver the widget for three weeks, she needs to be paid right now. The buyer does not trust the seller to keep her side of the bargain. He fears that if he pays in advance, the seller will deliver the widget late or not at all, and the buyer believes that any legal remedy would not provide sufficient compensation. One way that the seller might overcome the buyer's suspicion is to give the buyer a beloved pet as a hostage. If the seller fails to deliver or delivers late, the buyer can dispose of the pet. Because the buyer does not want the pet, he has no incentive to keep it rather than accept the widget; and because the seller values the pet more than the widget or the gains from late delivery, she will deliver on time. The

¹⁵ See THOMAS C. SCHELLING, *THE STRATEGY OF CONFLICT* 21-52 (1980) (discussing the way in which outside or self-imposed commitments affect bargaining relationships, strategies, and outcomes).

commitment mechanism in this case overcomes a barrier to cooperation.¹⁴

The two commitment mechanisms discussed so far—third-party deals and hostages—are not as important in social and commercial life as a third: reputation. If, in the first case, the buyer has a reputation as a hard bargainer, this could mean that people routinely charge him five rather than ten. The buyer can plausibly say to any particular seller, "If I paid you more than five, then everyone else would demand more than five, so I would lose a lot more than what I give you in this deal." If, in the second case, the seller has a reputation as an honorable dealer, she can say to the buyer, "I won't fail to deliver, because if I did, then my gains from this deal would be outweighed by my losses in every future deal that falls through because the buyer does not trust me to deliver." Reputation, like the contract to pay twenty to the third party and the hostages pet, enables its owner to commit himself to certain actions and therefore to overcome problems of opportunism in dealings with partners and competitors. This is why reputation is a valuable asset.¹⁵

Having a reputation means being "principled" in a certain sense.¹⁶ In ordinary talk, people often say that an unprincipled person acts out of self-interest and a principled person does not, instead conforming to the rules of morality or to the demands of altruism. This is not a helpful way of distinguishing between principled and unprincipled people. One often observes people who are considered to be principled violating their principles when the gains are sufficiently high. Moreover, the tough bargainer can be described as principled, but his principles serve his interests and have little to do with the requirements of moral theory. Finally, altruism is different from principled behavior. The parent who violates every rule of morality in order to help his child get ahead is an altruist, but is not a principled person. A person who makes charitable donations might be motivated by principle rather than by altruism. The tough negotiator is motivated by principle, not by altruism. The question is, then, how

¹⁴ See Oliver E. Williamson, *Credible Commitments: Using Hostages to Support Exchange*, 73 AM. ECON. REV. 519, 537 (1983) (arguing that "the use of hostages to support exchange is widespread and economically important").

¹⁵ See SCHELLING, *supra* note 13, at 35-43 (discussing the use of reputation to increase the value of a threat). The value of reputation as an asset is now a common theme in economics. See, e.g., DAVID M. KREPS, *A COURSE IN MICROECONOMIC THEORY* 531-36 (1990).

¹⁶ See SCHELLING, *supra* note 13, at 34 (analyzing the role of principles in bargaining).

can one explain principled behavior without abandoning the premise of utility-maximization.

One way to approach this question is to imagine that everyone is always in a relationship with someone else. Call this someone else a "partner," in order to abstract away from all the possible contexts: marital, social, commercial, and political. A relationship can be understood as a repeat prisoner's dilemma: In every round, each partner gains the highest payoff by "cheating" if the other partner "cooperates," the next highest payoff by cooperating if the other partner cooperates, the next highest payoff by cheating if the other partner cheats, and the lowest payoff by cooperating if the other partner cheats. As is well known, the rational move for each partner is to cheat if he does not care about future payoffs very much. If both partners care about the future sufficiently and the relationship has no definite ending point, there is a possible equilibrium in which both parties cooperate in every round.¹⁷

The reason that it is important that people care sufficiently about the future is that strategies that can maintain cooperation must threaten retaliation against anyone who cheats, and the only way to retaliate against someone who cheats is to deprive that person of future cooperation. Because most people care about the future, the threat of such retaliation can be effective. But it is reasonable to assume that different people value the future differently; in other words, people have different tastes about future payoffs ("discount rates"), just as people have different tastes for various goods. Young people, for example, are often impulsive: their decisions do not reflect much concern about the future consequences of their actions, so one says that they have high discount rates. Ambitious people who care deeply about earning fame or fortune later in life, parents who care about having enough money to support their children as they grow up, workers who want to have a comfortable retirement—all these people have low discount rates. Notice that, because discount rates are invisible, people can only infer the discount rate of another person, and they do so by observing his behavior.

In a cooperative relationship people with high discount rates cheat in circumstances in which people with low discount rates do not cheat. For example, some employees cannot resist sleeping late or skipping work even though they know that this behavior will reduce their future earnings; other employees are deterred from this misbe-

¹⁷ See, e.g., KREPS, *supra* note 15, at 503-15.

havior because they value their future earnings a great deal. Some wives cheat on their husbands even though they know that their infidelity might be revealed and, if so, will injure the long-term health of their marriage; other wives do not cheat on their husbands because they do not want to risk their marriage. Some suppliers deliver defective goods, because they do not expect to remain in business long; other suppliers deliver goods that are even above the stipulated quality or make other concessions to their buyers, because they want the buyers to continue doing business with them in the future. These patterns suggest a possible, but preliminary, definition of principled behavior. Because the people with low discount rates rarely cheat in the many relationships in which they take part, observers are likely to regard them as "principled." People who do cheat will be regarded as "unprincipled."

This definition is useful, because the explanation of principledness does not appeal tautologically to an assumption that people are motivated by principle, nor falsely to people's altruism, nor emptily to people's adherence to moral rules. Rather, the explanation appeals to the standard premises of rational choice theory. It also does not assume that people are "selfish": it takes no position on whether people's ends are self-regarding or other-regarding, or a mixture of both.¹⁸

The definition, however, has several defects. The main defect is that it does not explain the choppiness or discontinuity that characterizes principled behavior so much as the plain fact that some people cooperate more than others. Consider three people, *A*, *B*, and *C*. *A* has a very low discount rate, *B* has an ordinary discount rate, and *C*

¹⁸ To see this more clearly, note that two people, *X* and *Y*, who are both altruistic toward the other, face a kind of inverse prisoner's dilemma: each party wants to "cheat" by sacrificing the largest joint surplus in order to give the other party the higher payoff; but if both parties cheat, they end up at the jointly minimizing outcome. For example, suppose that each party has ten dollars, and can choose between giving the other party eight dollars at the cost of ten dollars (the loss of two dollars resulting from the cost of transacting), or keeping his ten dollars. *X* would think, "if *Y* keeps his ten dollars, then if I give *Y* my ten dollars, *Y* would be, with \$18, better off than if I keep my ten dollars; if *Y* gives me his ten dollars, then if I give *Y* my ten dollars, *Y* would be, with eight dollars, better off; therefore, I will give ten dollars to *Y*." But *Y* will think the same way, and they end up at the jointly minimizing outcome of eight dollars for each player. If they are principled and can resist the temptation to cheat (meaning they both keep their money), they end up with an outcome of ten dollars for each person. Cf. O. HENRY, *The Gift of the Magi*, in THE POCKET BOOK OF O. HENRY STORIES 1 (Harry Hansen ed., 1971) (providing an example where *X* and *Y* would have been better off if they had not exchanged gifts).

has a high discount rate. One might therefore observe that *A* never cheats in ordinary relationships, *B* rarely cheats, and *C* frequently cheats. The question is why people would say that *A* is principled and *C* is not, rather than simply saying *A* cares more about future payoffs than *C* does. Why do people tend to describe others in a dichotomous way—as principled or not—rather than saying that some people cooperate more than other people do? A further defect with the theory is that it does not explain self-presentation. Why is it that *A* would never say, “I have never cheated, but only because a sufficiently valuable opportunity has not yet presented itself”? Why would *B* and *C* insist that they are principled, and try to explain away their defections rather than admitting them?

One can solve these problems by relaxing two implicit assumptions of the model—namely, that people have perfect, or at least pretty good, information about the actions that other people have taken in the past, and that the only way to establish a good reputation is by cooperating in existing relationships. Suppose that *X* is trying to decide whether to enter a relationship (social, commercial, or political) with *Y*. *X* has no past experience with *Y*, so *X* cannot infer *Y*'s discount rate from his own observations of *Y*'s behavior. He knows that *Y* has a husband, several commercial partners, many friends, and he has heard that *Y* has never cheated on her husband or her friends, but that many of her commercial partners complain that she cuts corners. *X*'s problem is that he does not know how reliable any of the information—positive and negative—is. Maybe *Y* has cheated on her husband but is very discreet; maybe *Y* has not cheated on commercial partners but they are envious of her success and want to draw away her customers. If other desirable partners are available, *X* might refuse to enter a relationship with *Y*, despite her many attractions. By contrast, if *Y* has an unimpeachable reputation for integrity in all of her relationships, *X* will want to enter a relationship with her, too.

Because observers have imperfect information about a person's behavior in the past and because many people do not participate in many relationships over a long period of time, an agent's consistently cooperative behavior in a relationship may not suffice to establish a good reputation. The agent, therefore, has an incentive to engage in actions that will improve his reputation—actions in addition to, and other than, cooperating in existing relationships. Because an essential condition of a good reputation is a low discount rate, people try to establish good reputations by engaging in actions that show that they have a low discount rate. As noted before, discount rates likely are

distributed across the population: call those with low discount rates the "good" or "better" types, and those with high discount rates the "bad" or "worse" types.¹⁹ Because the better types care about the future, they can engage in costly actions as a way of demonstrating their concern for the future, as long as potential partners believe that only the better types engage in these actions, and thus are willing to enter long-term relationships with people who engage in these actions. If the action, or "signal," has the right cost structure, bad types cannot engage in it, because their (highly) discounted future payoffs are less than the cost of the action. However, there is a lot of noise in the world: different people interpret actions in different ways, they make different estimates about their costs, the costs of actions and payoffs from cooperation change over time, and so on. Accordingly, sometimes bad types are able to mimic the signals issued by good types; sometimes they will even take the initiative and issue new signals, with the result that the better types must scramble to mimic them.

As an example, one theory holds that firms invest in advertising in order to reveal that they are good types. Customers cannot observe the quality of goods prior to purchasing them, but they can observe quality afterwards. A good type, in this context, manufactures high-quality goods in the hope that customers will continue to do business with it and will tell other people about the firm. Before a seller has a reputation, it invests in advertising in order to show that it expects enough future business to recover the costs of advertising. Because bad-type sellers manufacture low-quality goods and therefore do not expect future business, they cannot afford expensive advertising. Even empty, meaningless advertising, so long as it is obviously expensive, operates as a signal that distinguishes the high sellers from the low sellers, so long as customers believe that any firm that spends a lot of money on advertising expects to obtain high returns in the future. In equilibrium, good sellers advertise and sell high-quality products at a high price, while bad sellers do not advertise (or advertise very little) and sell low-quality products at a low price, or do not make any sales at all.²⁰

¹⁹ I assume continuous types rather than dichotomous types; on this assumption and its problems, see B. Douglas Bernheim, *A Theory of Conformity*, 102 J. POL. ECON. 841, 846-47 (1994).

²⁰ See Benjamin Klein & Keith B. Leffler, *The Role of Market Forces in Assuring Contractual Performance*, 89 J. POL. ECON. 615, 630 (1981) ("[W]hen consumers do not know the minimum quality-guaranteeing price, the larger is a firm's brand name capital investment relative to sales, the more likely its price premium is sufficient to motivate high quality production.").

Individuals also send signals to each other in commercial and social settings. A common form of signal is the gift. One person donates a gift to another person as a way of showing a desire to maintain or enhance a relationship, or at least to forestall the inference that he wants to end the relationship. Gifts serve as signals because they are costly to the donor, either because they are expensive, or because they require an investment in information about the donee's tastes. In equilibrium, good types who have committed to a relationship send each other gifts, while bad types do not.²¹

But these forms of signaling will not always succeed, and they are likely not to succeed when signals are too noisy. Recall that in the advertising example the good types must charge high prices in order to recover the cost of manufacturing high-quality goods. A bad type offers a low price, claiming that the low price is due to more efficient methods of production, rather than to the low quality of its merchandise. The good types can distinguish themselves from the bad types only if customers believe that only good types can afford expensive advertising. Customers, however, might not believe that firms that purchase expensive advertisements produce better goods. The firms might try to implant this belief in the minds of customers by saying that their goods are better than those of their competitors, rather than using empty advertisements. But customers might not believe this statement, or they might not hear and remember it amid the cacophony of competing advertisements. Even if customers believe that firms using expensive advertisements are more likely to produce better goods, they might not be able to distinguish expensive and not-so-expensive advertisements. In such a case, a bad firm that purchases not-so-expensive advertisements is indistinguishable from a good firm, and the good firm will mimic the bad firm's advertising strategy in order to save money.

In equilibrium, both the good and the bad firm will say that they always deliver high-quality goods in their not-so-expensive advertisements. There are two theories for why they will make this statement. First, they must say *something* in order to persuade buyers that the advertisement shows that they are good types, and the natural thing to say is that they are good types. They will say "I never cheat" (that is,

²¹ See Colin Camerer, *Gifts as Economic Signals and Social Symbols*, AM. J. SOC. S180, (Supp. 1988) (analyzing conditions under which gifts can serve as signals); Eric A. Posner, *Altruism, Status, and Trust in the Law of Gifts and Gratuitous Promises*, 1997 WIS. L. REV. 567, 579-80 (discussing how gift-giving distinguishes cooperators from opportunists).

“my goods are high quality”), rather than “I usually never cheat” or “I cheat one percent of the time,” because the claim that one never cheats is *focal* in Schelling’s sense.²² A firm that wants to persuade buyers that it is reliable makes an absolute claim of reliability (100% reliable), rather than a partial claim (98% or 97.5% reliable), even if the partial claim is more honest, because people do not differentiate among shades of reliability and will interpret anything less than a claim of absolute reliability as an admission of substantial unreliability. Given that in signaling models *any* action with the right cost structure can be a signal, the most natural statement in the advertising context is a claim that the goods are always high-quality.

The second theory for why all kinds of firms will claim never to cheat is that some buyers might be credulous (by which I mean poorly informed), even if most buyers are not. If these buyers accept statements at face value, then they will buy from the firm that makes the strongest claim that it does not cheat. Any firm that honestly says that it cheats 10% of the time will lose these buyers to a firm that says that it cheats 9% of the time. Firms will bid each other down to claims of 0% cheating in order to avoid losing buyers to each other.²³ The same result is obtained by assuming that none of the buyers is poorly informed but that a tiny fraction of firms really is principled.

To be clear about why people act this way, put yourself in the shoes of the buyer in the second example. If the seller tells you that she never cheats, then you might not believe her, but because all the other sellers say the same thing, this statement gives you no basis for rejecting this seller in favor of some other seller. But if the seller tells you that she sometimes cheats, while all other sellers deny ever cheating, you might as well take your chances with another seller. One might argue that the buyer should admire the seller who admits that she cheats: this seller is, at least, honest. But someone who cheats is, by definition, dishonest. If she says that she cheats, you have no reason to believe that she is more honest than people who deny that they cheat. It is at least possible that someone in the group of people who deny that they cheat really is honest; it is impossible that the self-proclaimed cheater is honest.

Even if you are so cynical that you would never believe absolute or extreme claims, you are not likely to treat differently the seller who makes absolute claims and the seller who does not. You gain nothing

²² See SCHELLING, *supra* note 13, at 57 (introducing the idea of focal points).

²³ It is possible that, instead, markets will segment.

by treating them differently. (Indeed, if you and others tried to treat them differently, they would conform their behavior to your expectations without being any less likely to cheat you.) But if at least some people are credulous, sellers benefit by making absolute claims. These claims cost them nothing. They do not lose the cynics' business, and they gain, or avoid losing, the business of the credulous people.

In sum, firms are forced to make exaggerated claims about the purity of their motives and the quality of their products. These exaggerated claims of never cheating can be described as claims to be principled, while the claim that one occasionally cheats is not a claim that one is principled.²⁴

The analysis applies not only to relationships between firms and consumers. It applies to any cooperative relationship. Everyone will claim to be principled as well as invest resources in advertisements or gifts. The bad or worse types will claim to be principled either because of the possibility that someone will believe them or because principledness is focal. The good types have to follow the bad types or risk being mistaken as bad types. Returning to the earlier examples, the buyer will say, "I never pay more than rock-bottom price." He will never say, "I never pay more than rock-bottom price unless I'm really desperate for the goods," even though the latter statement is surely more accurate than the former. In the second example, the seller will say, "I never deliver late," rather than, "I deliver late only when I am having serious labor problems." In social relationships, people say to their friends and lovers, "I will never betray you," or "you can count on me," rather than, "you can count on me unless my gains from the relationship fall below a certain threshold."

I am not arguing that a seller would never say that it will not have labor problems that might interfere with its ability to perform. Some sellers do say that, but that is not the same as saying that one will cheat. When the seller informs the buyer of the possibility of labor problems interfering with delivery and the buyer nonetheless accepts the seller's offer—no doubt with a price adjustment that reflects this

²⁴ Technically, the claim that one is principled is not a signal because it is not costly. I have treated it as "outside the signaling model," simply as people's description of what they are doing "in the model." A more precise and complete model would treat the claim that one is principled as "cheap talk," but the same result would obtain in a cheap-talk model. See, e.g., Stephen Morris, *An Instrumental Theory of Political Correctness* (Jan. 1998) (unpublished manuscript, on file with the *University of Pennsylvania Law Review*) (showing that when reputation matters enough, an informed individual cannot convey certain kinds of information to an uninformed individual).

risk—then failure to deliver because of labor problems is not cheating. The value-maximizing obligations that emerge in a relationship can be rigid or flexible, depending on context.

Two further points should be made. First, a person who claims to be principled is not being honest. As mentioned above, a rational person will sacrifice his reputation when the gains are sufficiently high. But a principled person is one who claims never to make such a tradeoff. The person who claims to be principled does not say: "I will not cheat unless the gains exceed reputational losses." He says: "I will never cheat." People are driven to extreme statements by their focal nature and by the need to match the exaggerated claims of bad types. Thus, in equilibrium all people claim to be principled, but they are believed only by credulous people (such as children, whose discovery of this fact about the adult world is the beginning of their disillusionment). This is why an air of corruption hangs over every bargaining table.

Second, good types try to distinguish themselves from bad types by behaving consistently with their words as much as possible and by making claims that can be verified. The claim that one is principled cannot be disconfirmed by the addressee until it is too late—until the speaker has cheated. But the speaker might overcome the cheapness-of-talk problem by integrating claims about his principledness into a general theory about himself. Consider the following common statements: "I feel terribly guilty when I tell a lie." "My parents taught me never to take advantage of other people." "God punishes those who tell lies." "In the army we were taught always to tell the truth." Statements like these are more plausible than acontextual assertions of truthfulness because the speaker reveals other aspects about himself that can then be used to verify his truthfulness. If the first speaker later plays a nasty practical joke, the second speaker turns out to be an orphan, the third an atheist, and the fourth a draft dodger, then the listener has reason to doubt the speaker's claim to be a truth-teller.

Thus, in their effort to distinguish themselves from bad types, some good types engage in speech and behavior in which they would not engage but for the problem of distinguishing themselves. Note that good types are driven not only to distort their speech, but also to distort their behavior. I might claim that I am principled because I spend long hours at church. If spending long hours at church is sufficiently costly and differentially costly, I might distinguish myself as a good type by making this claim, but only if my claim can be verified.

If others observe that I do not go to church, I will be exposed as a bad type. I go to church, then, as a way of showing that I am principled, not because I seek religious solace. My investment in this costly signal enables partners to cooperate with me to a greater extent than with bad types who cannot afford the signal. In this sense, the signal serves to enhance cooperation. Still, I am not *principled*, because I will cheat if an exogenous change modifies my payoffs by a sufficient amount.

Before moving on to the issue of incommensurability, let me summarize the argument so far, point out one additional implication, and respond to an important objection. The argument is that people acting in their rational self-interest face powerful incentives to claim to be principled in the ordinary sense of the word; that is, never to cheat in the cooperative relationships to which they belong. Having made the claim to be principled, people must act fairly consistently with the demands of principle, lest they be immediately exposed as bad types. At the same time, people will cheat when the short-term gains outweigh the long-term reputational costs. Given the focal character of absolute statements, and the existence of some people who are credulous, one has a strong incentive not to admit the truth, and to claim, to the contrary, that one never cheats. Since, in equilibrium, no person admits that he would ever cheat, the statement that one does cheat would cause people to infer that one belongs to the bad type.

It is worth emphasizing that because sometimes a person's principled claims will constrain his behavior, the person will cheat less often (though he will cheat if the payoff is high enough) than he would if he did not make principled claims. Therefore, claims to be principled actually may produce social benefits—by reducing the amount of cheating—even though they do not ensure or reflect principled behavior.

The objection to the theory is that it treats people like “rational fools,” in Sen's words.²⁵ Sen argued influentially that economics assumes that people are rational calculators, whereas in fact people are often motivated by principle, or what he calls “commitment.”²⁶ My theory shows why people act *as if* they were motivated by principle, and it therefore addresses Sen's criticism to the extent it is intended as a critique of the explanatory ambitions of rational choice theory.

²⁵ Amartya K. Sen, *Rational Fools: A Critique of the Behavioral Foundations of Economic Theory*, 6 PHIL. & PUB. AFF. 317 (1977).

²⁶ *Id.* at 317-18.

But Sen also argued that the methodology of economics misses an important side of human motivation, phenomenologically speaking.²⁷ The problem with Sen's argument, however, is that simply assuming that people act out of principle *and* rational calculation gives one less methodological purchase than the ordinary rational choice assumptions do, without, as far as I can tell, compensating for this loss by producing a methodological gain.²⁸ One effort to solve this problem without abandoning the rational choice perspective is Frank's theory of the motivational force of emotions. According to Frank, people retaliate at appropriate times and decline to cheat because they are compelled by their emotions.²⁹ But Frank does not explain why people sometimes can overcome their emotions and sometimes cannot, with the result that in his theory, principles operate as a kind of drag on rational calculation but do not substitute for it.

In any event, I concede, following Sen and Frank, that the theory I present is not supposed to represent human motivation in all its complexity. Instead, it abstracts out one element of human motivation in the hope of shedding light on the problem with which I began. After all, this problem—why people represent themselves as principled when they are not really—has not been solved by people who take alternative methodological approaches.

II. INCOMMENSURABILITY

I have used the idea of being principled in a broad sense, pretty much as the equivalent of being a person who does not cheat in relationships, a cooperator. People claim to be principled because others infer that people who do not make such claims are bad types. Similarly, people claim that options are incommensurable because others infer that people who do not make incommensurability claims are bad types.

To understand this point more clearly, suppose that people believe most things a person says about himself, and this person wants people to enter cooperative relationships with him. What sort of things might that person say? Recall that a person is an attractive co-

²⁷ See *id.* at 317.

²⁸ Cf. JON ELSTER, *THE CEMENT OF SOCIETY* 35-37 (1989) (describing social norms, altruism, envy and spite, and rational, selfish process- and outcome-oriented motivations). Elster's description of the riot of human motivations provides a phenomenologically sound, but methodologically sterile, basis for understanding social behavior.

²⁹ See ROBERT H. FRANK, *PASSIONS WITHIN REASON* 53 (1988).

operative partner if he has a low discount rate, because a person with a low discount rate is relatively unlikely to cheat in a game characterized by indefinitely repeated transactions. But the low discount rate is not a sufficient condition for cooperation. It must also be the case that the person gains, in any given round, relatively little by cheating when the other person cooperates. So, in order to attract cooperative partners, a person will say that (1) he has a low discount rate and (2) his outside opportunities are not valuable.

People make these claims in complex ways. First, they claim to be mature rather than immature, self-disciplined rather than impulsive, strong rather than weak. All of these claims are about discount rates. Second, people claim that if they cheated they would feel guilty or ashamed, they would burn in hell, or they would not be trusted by other people. These claims all are about outside opportunities. Third, people claim that they are principled or honorable, or that they do not engage in rational or strategic behavior. These claims constitute a denial that one is even capable of engaging in the sort of calculations that make strategic behavior a danger in the first place.

When a person makes an incommensurability claim, he is making a claim of either the second or the third kind. One incommensurability claim is that the benefits from cooperating cannot be compared with the benefits from cheating. To admit that the two kinds of benefits are comparable would be to say that one will choose the action (cooperation or cheating) that provides the higher benefit. Denying the possibility of comparison is a denial that the speaker can engage in strategic behavior and, by implication, a guarantee that the speaker can only cooperate. The incommensurability claim is the claim that there is no question of a tradeoff: The benefit from cooperation is "differently valuable"³⁰ from the monetary gains from cheating.

The other incommensurability claim is that the benefits from cooperating are infinitely high. If the benefits from cooperation are infinitely high, whereas (by implication) the benefits from cheating are finite (if sometimes very high), even a rational maximizer will never cheat, if that means that in a later round the gains from cooperation will be lost. The incommensurability claim is again that there is no question of tradeoff, but this time because the benefits of cooperation are so high.

³⁰ See Sunstein, *supra* note 1, at 813.

These two ways of claiming to be principled track a dispute between Sunstein and Regan. Regan argues that the incommensurability claim implies that the speaker places an infinite value on something.³¹ Sunstein argues that the incommensurability claim implies that the speaker places a different kind of value on something.³² Because people make both kinds of claims, and both kinds of claims serve the same general strategic purpose, for my argument there is no need to make the distinction.

As an example, consider a person who refuses to answer a cost-benefit survey about the value of a mountain view, claiming that no amount of money can compensate him for obstruction of that view. In this person's community, the commitment to the environment serves as a signal of one's loyalty to others; anyone who says that the mountain view can be traded for something else would be considered a bad type. To show that he is not a bad type, this person would deny the possibility of such a tradeoff. This claim could be interpreted as an assertion that the mountain view has an infinite value, or as an assertion that comparisons cannot be made—it does not matter. Both claims serve the strategic purpose of signaling one's loyalty to the relevant community.

Why is the simpler interpretation not that the person places an infinite or incommensurable value on the mountain view? The answer is that, except in extremely rare circumstances, people do not act as if they placed an infinite value on anything: Individuals routinely make tradeoffs in their lives. For example, people sacrifice mountain views to move to a location where there are better jobs, better schools for their children, and more culture. When pressed, people admit that they would give up a mountain view if that meant saving lives, producing medicines, or reducing crime and unemployment. A person who places an infinite or incommensurable value on something *in fact* would give up everything he owned, every future opportunity, every good for other people, in order to maintain that invaluable thing. If there are any people like this, they are too few to have strong claims on the direction of public policy.

A response to this argument is that the person would not trade the mountain view for money, but would trade it for other goods, such as reduced unemployment in the region. That is, the mountain

³¹ See Donald H. Regan, *Authority and Value: Reflections on Raz's Morality of Freedom*, 62 S. CAL. L. REV. 995, 1058 (1989).

³² See Sunstein, *supra* note 1, at 813.

view and money are incommensurable, but the view and employment are comparable. But it is hard to see why, if a view and employment are comparable, they cannot both be compared to money, which serves as a metric. A sum of money has as much in common with a mountain view and with a reduction in unemployment, as the latter two have with each other. The implicit assumption here is that the sum of money would be used for a self-interested, rather than an other-regarding, purpose. *That* is why trading the mountain view for money would be objectionable, just as trading the mountain view for a private swimming pool would be objectionable. A person in the hypothetical community could *overtly* make a tradeoff between the mountain view and the reduction in unemployment, but not between either of these options and a sum of money or a private swimming pool, because one obtains reputational gains by professing respect for the quality of life of others, not by demonstrating concern about one's possessions. People who care about themselves too much are not as attractive cooperative partners as—are more likely to cheat than—people who care about others as well, because we do not think that people are as likely to violate obligations for the sake of other people, or for abstract classes of other people, as they are for their own sake. (People who are too altruistic, however, are avoided just as vigorously as people who are too greedy; both kinds of people are unreliable.)

Another example is a professor who refuses to accept money to skip a faculty meeting but who would skip the faculty meeting in order to stay home with a sick child when child care is not available. One might believe that these three options—going to the meeting, taking the money, or staying with the child—are incommensurable. Preliminarily, note that the monetary option is unacceptable not because it involves money. Most people would think that the professor could skip the faculty meeting if the government threatens to fine him \$100,000 if he attends. The reason that the bribe is unacceptable and the threat and the child care option are acceptable is that the implicit, value-maximizing obligations that emerge among faculty members allow one to skip meetings in order to meet serious personal obligations that conflict with them, or to avoid serious harms. These personal obligations do not allow one to satisfy predictable fluctuations in tastes or to seize predictable opportunities. This view is standard in contract theory, and it tracks Raz's theory of exclusionary reasons, which is related both to his argument about incommen-

surability and to a discussion of promises.³³ Still, nothing about the example suggests that such behavior is inconsistent with the choices of rational actors.

A related example is the person who refuses to trade time with his family for money, but would trade time with his family for a better career opportunity. This person does not claim that time with his family has infinite value. Rather, he claims that it cannot be compared with money or even with the better career opportunity on a common metric. It is hard, however, to see what is accomplished by this argument. Trading time with one's family for money would likely count as cheating on the family relationship. Trading time with one's family for a better job opportunity would count as cheating in some families, but not in others—it simply depends on what value-maximizing obligations emerge over time. Incommensurability arises because people promise not to trade time with the family for anything. This absolutist claim is, again, more focal and more resistant to the competing exaggerations of bad types than a more honest and limited claim, such as, "I will not trade off time with the family except for something that is very important to me." The moral awkwardness of the tradeoff reflects the sense that one is caught in a game that denies its players the luxury of honesty.

As discussed above, this is not to say that incommensurability claims are empty talk. In fact, the moral quality of these claims puts people in a bind. On the one hand, one cannot truly act consistently with them. On the other hand, if one does not act consistently with them, one may lose the reputation that the claims are intended to establish. People shade their behavior so that violations of their incommensurability claims occur relatively infrequently in ordinary life, but they do not act consistently with them. The problem torments politicians because their words and actions are scrutinized so carefully by the press. The politician who refuses to make incommensurability claims will be called a pragmatist by his friends, and a person lacking in vision by his enemies. The politician who makes incommensurability claims will be called principled or rigid. When the latter is forced to compromise, he is then called slick. Ordinary people can avoid these problems as long as the incommensurability claim is consistent with rational action, that is, when cooperation remains rational; when not, they cheat and try to conceal their opportunism behind casuistry.

³³ See JOSEPH RAZ, *PRACTICAL REASON AND NORMS* 69-71 (1990) (drawing an analogy between the formal features of promises and decisions as reasons for acting).

A third example is the friend who cancels lunch. A person *X* who offers his friend *Y* twenty dollars as compensation for a canceled lunch sends the wrong signal—it violates the “ideology” of friendship. (The loaded term “ideology” will make a reappearance later in this paper.) According to the ideology of friendship, a friend never cheats another friend. *Y* might think that if *X* offers twenty dollars, then *X* must believe that *Y* thinks that *X*’s cancellation was a form of cheating that would cost *Y* twenty dollars. If *X* thinks that *Y* places a monetary value on this (minor) form of cheating, then *X* must think that *Y* would also place a monetary value on more significant forms of cheating or betrayal. If *X* thinks that *Y* can be compensated for any betrayal, then *X* must think that *Y* puts a monetary (and finite) value on their friendship. That implies, however, that *X* thinks that *Y* would cheat if the payoff were sufficiently great. By offering twenty dollars, *X* implies to *Y* that he believes that *Y* is not really a friend, which, of course, offends *Y* as much as if *X* said to *Y*, “you’re a liar.”

But why doesn’t *Y* say to herself, “*X* is no saint, but he still can be my friend; only saints place an infinite value on friendships”? The answer is that absolutist claims are focal and resistant to the exaggerated claims of bad types. If *X* values *Y*’s friendship at $\$A$ and reveals this to *Y*, then someone else will claim that he values *Y*’s friendship at $\$A + 1$; then another person will claim that he values *Y*’s friendship at $\$A + 2$; and so on. Only an infinite valuation cannot be trumped by an effort by a bad type to insist that he values the friendship so much more. What frustrates the good type is that, although by claiming to value the relationship at an infinite amount he avoids the inference that he belongs to the bad type, he also is unable to distinguish himself from the bad types and he also has been corrupted by his emulation of them.³⁴

Incommensurability claims occur because money is a problem. The problem with money is that, by supplying a scale for ranking all options, it exposes the strategic basis of cooperation, and, by exposing the strategic basis of cooperation, it interferes with cooperation. Given the observable alternatives supplied by the market, the value a person places on a particular relationship can be estimated within a range. That person must exaggerate the value he places on the relationship in order to avoid the inference that he belongs to the bad type and persuade his partner that he values the relationship at the

³⁴ Cf. WILLIAM SHAKESPEARE, *KING LEAR* act 1, sc. 1. To be sure, Cordelia opted out of Goneril and Regan’s bidding war; but look at what happened to her.

upper end of the range. The exaggeration is an embarrassment, however, because everyone knows that it is an exaggeration. This is the source both of the stigma that is associated with money and of the market, and the feeling of alienation that the market economy is said to produce.³⁵

The stigma reflects the same dynamics that give rise to incommensurability claims: People stigmatize others who too overtly desire money in order to represent, by comparison, their own immunity from being bought off or tempted to defect by a higher payoff. This is why people must constantly deny that they are motivated by money even as they rush around accumulating it. The feeling of alienation is the result of the dissonance between one's behavior, which inevitably reflects tradeoffs, and the public denial of this fact.

The analysis is the same for neighborly, intimate, work, and political relationships. These are all cooperative ventures in which it is important to signal that one is not a cheater. Parties to these relationships continuously send each other signals of their loyalty, precisely because they know—and they know that their partners know—that it may be worthwhile for them to break off the relationship. One way to signal loyalty is to say that the relationship is incommensurable with other options. Few people really believe this, but no one can express their skepticism openly without risking the end of the relationship.

As we saw earlier, the claim that environmental goods are of infinite or incommensurable value also illustrates the use of incommensurability claims to signal that one will not cheat. People who are in environmental groups (defined broadly to include people who are in any type of relationship and who also value the environment a great deal) show their loyalty to other members by saying that they place infinite or incommensurable value on the goals to which the group is devoted. People do this in order to persuade others that they do not belong to the bad type. This logic can be extended to other activities. Teachers tell each other that teaching is of great value in order to persuade each other that they are loyal to the group endeavor. So do research scholars, doctors, lawyers, actors, artists, and, no doubt, everyone else. Telling a coworker that "we are paid what we are worth" is not going to endear you to him; telling him that "we are all under-

³⁵ Cf. SIMMEL, *supra* note 1 (discussing the alienating effect of money); cf. also NIGEL DODD, *THE SOCIOLOGY OF MONEY* 46-52 (1994) (outlining Simmel's discussion of the alienating effect of money).

paid, and no one appreciates the value of what we do" might. The person who makes the first statement cannot be trusted in, say, the formation of a union; the person who makes the second statement at least *might* be trustworthy.

An objection to the argument so far is that people are not as cynical as the argument implies. The argument, however, does not claim that people are cynical, or that they make incommensurability claims cynically. Many people believe their incommensurability claims. One can sincerely make such claims as long as circumstances do not provide a valuable opportunity to cheat. As long as the demands of interest and of ideology converge, there is no cognitive dissonance that must be resolved, at least for an unreflective person. When they do conflict, the usual response is to rationalize one's defection; that is, to revise one's story about oneself or about the action that is believable and that preserves one's reputation or self-image as a principled person. For example: "That was not really cheating, because the other person was about to cheat me, anyway"; or, "Everyone has a lapse now and then, and this won't happen again."⁵⁶

III. IMPLICATIONS

It would be convenient if the theory I have described implied that the government should simply ignore incommensurability claims when determining policy, but life is more complicated. Imagine that the government must decide whether to preserve a small patch of forest or to allow companies to exploit it. The benefit from exploitation is the creation of jobs for impoverished members of a Native American tribe and the production of medicine that would reduce the suffering of AIDS patients. The cost is the elimination of a piece of the environment, which, according to surveys of some nearby residents and some citizens living elsewhere in the country, has infinite or incomparable value. What should the government do?

The simple response is to say that the government should either ignore the infinite valuations or should estimate the amount of money that the protestors would trade off in order to preserve the piece of forest. Because these people do not really value the project at negative infinity and do not really believe that the losses and the

⁵⁶ See the discussion of "avoidance" in LEO KATZ, *ILL-GOTTEN GAINS: EVASION, BLACKMAIL, FRAUD, AND KINDRED PUZZLES OF THE LAW* (1996) (outlining the relationship between "sleazy evasion" and "circumspect avoidance" of legal and moral rules).

gains are incommensurable, their survey answers should carry little weight with the government.

The first problem with the simple response is that it ignores the important role that incommensurability claims play in social, commercial, and political life. Even though incommensurability claims do not reflect true valuations, they are often valuable as forms of behavior. They help maintain cooperative relationships, and cooperative relationships are valuable as long as they do not injure third parties.³⁷ If the state could forbid people to make incommensurability claims, this prohibition could (though not necessarily) make people worse off, because it would be harder for people to signal their suitability for a cooperative relationship. Of course, the state could not forbid people to make such claims; but even ignoring such claims might disvalue them in a way that reduces their utility for strategic purposes. Indeed, from a utilitarian perspective, the incommensurability problem puts the government in a bind. If the government takes incommensurability claims seriously, it must fail to make trade-offs that increase social welfare. But if it discounts them, it might interfere with people's efforts to engage in valuable cooperative ventures.

Still, one should distinguish between the existence of incommensurability and other principled claims, and what these claims are *about*. The government cannot prevent people from making *any* incommensurability claim, and even if it could, there is no reason to believe that the new equilibrium would be characterized by greater social welfare. But some incommensurability claims are more socially costly than others. People for whom ethnic purity is a matter of principle might be superb cooperators within their own ethnic group. These people claim that they will not cheat on members of their own group in order to obtain benefits from outsiders. This is an incommensurability claim because they deny that any amount of outside benefit will break the bonds of ethnic loyalty. The effectiveness of this claim at promoting cooperation can be seen in the history of nationalist movements, most recently in the former Yugoslavia.³⁸ In the

³⁷ If people can inculcate themselves or their children with a commitment to incommensurability, they might make themselves into more reliable cooperative partners. Cf. FRANK, *supra* note 29 (proposing a similar role for the emotions); Frederick Schauer, *Instrumental Commensurability*, 146 U. PA. L. REV. 1215 (1998) (making an analogous argument at the level of social choice).

³⁸ See RUSSELL HARDIN, *ONE FOR ALL: THE LOGIC OF GROUP CONFLICT* 156-63 (1995) (discussing the strategic structure of conflict in Yugoslavia).

United States, this claim frequently has been advanced by whites as a way of strengthening their bonds with each other. Yet clearly this kind of incommensurability claim deserves no respect from the government, for the advancement of this claim commits the speaker to harming third parties. A similar argument can be made about religious incommensurability claims. By comparison, incommensurability claims about one's family and friends are relatively, though not completely, benign.

Incommensurability claims about the environment present similar difficulties. Environmental protection is a principle around which people rally. They make incommensurability claims about the environment in order to show their commitment to like-minded people. Environmental incommensurability claims, unlike ethnic incommensurability claims, do not commit the speaker to behavior that directly harms third parties. The problem with these claims is that, if acted on, they will indirectly injure third parties—people will be denied jobs, medicine, and inexpensive food because of restrictions on development. I am not arguing that development is necessarily good—I am arguing that incommensurability claims interfere with efforts to evaluate development according to its contribution to the welfare of those affected, and that this interference is not justified by a vindication of authentic moral values. The interference is, rather, a side effect of the fact that certain political positions have become symbolic of a person's trustworthiness within certain groups.³⁹ Similar comments can be made about incommensurability claims concerning bodily integrity, including the claims that people should not be permitted to sell their organs, or that women should not be permitted to bind themselves to surrogacy contracts, whatever the benefit to them.⁴⁰

The second problem with the simple response—that the government should simply ignore or discount incommensurability claims—is that it is paternalistic. On what grounds should the government ignore the expressed views of its citizens? The answer is that the government should ignore these views when they do not reflect citizens'

³⁹ See Eric A. Posner, *Symbols, Signals, and Social Norms in Politics and the Law*, 27 J. LEGAL STUD. (forthcoming 1998) (analyzing symbolic behavior).

⁴⁰ See Margaret Jane Radin, *Market-Inalienability*, 100 HARV. L. REV. 1849, 1937 (1987) (concluding that it is incompatible with "human flourishing" and "our conception of personhood" to allow some things to be objects of exchange in the marketplace). Although Radin prefers the term "commodification," this term and "incommensurability" are used interchangeably in the literature.

interests and values. When citizens make incommensurability claims in order to avoid reputational penalties, preferences will be concealed, and preference-respecting policies would not reflect citizens' true preferences.⁴¹ Under these circumstances the government should not defer to expressed preferences.

This is a theory of ideology. An ideology consists of propositions that people publicly endorse in order to show that they are good types. Public endorsements of these propositions count as signals of loyalty because people infer that those who reject these propositions are bad types. Ideologies are exaggerated and simplified versions of the world views that people actually hold. The nuance and blurriness of people's nonideological views about the world disappear in the ideological version because nuanced positions do not serve well as signals, whereas extreme positions do. Ideologies are rigid: they rise suddenly and collapse violently. World views are flexible: they shift and evolve gradually, often imperceptibly. The rigidity of ideologies and the discontinuity of ideological change result from the self-reinforcing nature of signaling equilibria. As long as everyone takes the public endorsement of certain positions as evidence of loyalty, people will be reluctant to disagree; but once changes in circumstances increase or decrease the cost of deviation by a sufficient amount, the equilibrium collapses.⁴²

This theory can be usefully contrasted with deliberative theories of preference formation. These theories argue that preferences shift in valuable ways when exposed to the winds of public debate.⁴³ This claim is no doubt true, in part. It must also be acknowledged, however, that preferences can harden when emotions run high; people take refuge in whatever view emerges as dominant in order to avoid reputational penalties. People conceal their actual interests and values because of the importance of signaling, and the signals them-

⁴¹ Cf. TIMUR KURAN, *PRIVATE TRUTHS, PUBLIC LIES* 105-17 (1995) (arguing that "preference falsification is a...reason for the persistence of unwanted social choices"); Glenn C. Loury, *Self-Censorship in Public Discourse*, 6 *RATIONALITY & SOC'Y* 428, 430 (1994) (arguing that the "political correctness" movement has led to "self-censorship" on the part of those who are afraid of appearing to violate community norms).

⁴² See KURAN, *supra* note 41, at 247-60 (explaining how rapid "unanticipated revolution[s]" are possible "[w]here the status quo owes its stability to preference falsification").

⁴³ See, e.g., CASS R. SUNSTEIN, *FREE MARKETS AND SOCIAL JUSTICE* 20 (1997) (describing the role of deliberative democracy in the formation of preferences).

selves are distorted and made extreme by the focal character of absolute statements.

The role of the critic is to identify those popular propositions that are arbitrary focal points of loyalty and that do not reflect people's authentic interests. Critical theory seeks to identify distortions in the political process that undermine political representation. One such distortion is reflected in the equilibrium that results because people have strong incentives to signal their desires to enter and enhance relationships with each other. Because people do not want to offend their partners, they will not make certain criticisms which, if followed, would make everyone or most people better off. A purpose of critical theory, then, is to describe conditions under which such an equilibrium results, and to criticize the statements and actions that serve as signals in that equilibrium. Habermas's proposal that we imagine an ideal conversation in which people are forced to provide honest defenses to any claim they advance is an approach to this problem.⁴⁴ In such a conversation, one would have to admit that one supports a particular policy in order to signal one's loyalty; thus, one's true preferences would be revealed.

The use of signaling models to disclose ideological distortions in public debate falls short of the ambitions of Habermas and other critical theorists. It would not show that any nonreputational preferences are ethically objectionable. But so ambitious a version of critical theory runs up against serious conceptual barriers.⁴⁵ The marriage of signaling theory and critical theory avoids these problems while also having a great deal of purchase. Such a theory would allow us to identify areas of life and politics in which the perceived importance of signaling one's type overwhelms the incentives to speak or act truthfully, resulting in equilibria in which nonreputational preferences are not satisfied. The critical theorist would have the role of persuading people that they would be better off if they acted and believed differently: his or her "imaginative reconstruction" would propose a new focal point around which behavior would shift.⁴⁶ Although I cannot go into detail here, a critical theory along the lines I have suggested

⁴⁴ See JÜRGEN HABERMAS, *THE THEORY OF COMMUNICATIVE ACTION* (1984).

⁴⁵ See, e.g., RAYMOND GEUSS, *THE IDEA OF A CRITICAL THEORY* (1981).

⁴⁶ As noted before, reputational concerns distort not only people's claims, but also people's behavior. This suggests that economists should be skeptical not only of survey data (a skepticism that is frequently expressed in the literature), but also of market behavior, when trying to determine preferences.

could expose the focal character of propositions underlying such ideologies as consumerism, religious fundamentalism, and nationalism.

This leads to the third problem with the simple response that incommensurability claims should be ignored by the government. There is a vast gulf between the normative project of identifying distortions in people's political views and the political project of forcing people to change their views. Incommensurability claims present a political problem. On the one hand, the government cannot take them seriously, because then it could not make the tradeoffs that governments have to make all the time. On the other hand, if the government does not take them seriously, people will press their incommensurability claims against the government—again, as a way of showing each other that they belong to the good type. This conflict is played out over and over on the stage of American politics. Economists, technocrats, and pragmatists trade off, balance, and compromise. Religious figures, moral leaders, and ideologues respond by asserting the impossibility of compromise over matters of principle. Politicians, in turn, are forced to take a position on these issues, as a way of signaling that they, too, are principled. American politics mirrors the conflicts faced by ordinary people: on the surface we see furious controversies over flag desecration, abortion, gun control, affirmative action, and other issues on which all politicians are required to have a firm and principled position. Beneath the surface we see frequent compromise—by the same people—over pork, taxes, and intricate regulatory statutes that lack the national salience of the other issues.

If you look now at Sunstein's list of the implications of incommensurability for various legal rules and problems, the *conceptual* problems that he identifies become less serious.⁴⁷ People who take absolute positions on religion, free speech, and environmental protection are signaling to relevant communities that they can be trusted. They take absolute positions, rather than merely strong positions, because they fear being "outbid" by bad types who claim to be more deeply committed than they are, and because extreme positions are more focal than nuanced views.

However, the *political* problem remains. People will continue making incommensurability claims because of the pressure to signal trustworthiness. People will resist attempts by the government to discount incommensurability claims; indeed, politicians and other gov-

⁴⁷ See *supra* notes 5-10.

ernment officials will not want to discount them too much, lest people infer that they are bad types. Still, the political problem should be distinguished from the normative problem. We care about incommensurability claims because they inevitably emerge in social life, but we should not confuse them with deeper values.

CONCLUSION

This Article has not refuted the arguments advanced by philosophers who believe in the incommensurability of values. Instead, it has shown that the evidence on which they base their arguments—including certain statements people make, certain choices they make, and the moral uneasiness they feel about these choices—is susceptible to another interpretation. On this interpretation, incommensurability claims and other principled claims emerge in an equilibrium in which rational actors try to show that they are unlikely to cheat in cooperative endeavors.

This interpretation, if correct, has an important implication: A government that maximizes social welfare should discount incommensurability claims made by citizens, unless such discounting would interfere with socially valuable cooperation.⁴⁸ More generally, the welfare-maximizing government should treat with skepticism any valuations—whether reflected in words or in market behavior—that reflect ideological commitments, where by “ideology” I mean the cluster of propositions public endorsement of which serves as a signal of group loyalty.

The theory presented in this paper says nothing about self-conception, and perhaps it hollows people out too much. There is a psychological residuum not explained by the essentially behaviorist theory that I have described. Objects, as the old idea has it, have both an intrinsic value and a value that represents their exchangeability for other, more desirable objects. Thus, we enjoy our houses but with half a mind toward whether we can trade them up for better houses. Relationships are similar. One enjoys one’s relationships for their intrinsic value but also knows that they can be traded for even better relationships: this is what is meant by betrayal. This knowledge makes us uneasy, because we know that our partners may betray us, and we

⁴⁸ Cf. Richard Craswell, *Incommensurability, Welfare Economics, and the Law*, 146 U. PA. L. REV. 1419 (1998) (assuming that “theories of incommensurability are correct as far as individual decisions are concerned,” but finding that as applied to governmental decisions, those theories are untenable).

know that *they* know that *we* may betray them. Our trust and distrust drive us endlessly to provide evidence of our good faith, and to regard theirs with skepticism. No doubt this tension existed in societies that did not have money, but the invention and spread of money must have increased the tension, simply because money—"the most certain image and the clearest embodiment of the formula of all being, according to which things receive their meaning through each other," in Simmel's words⁴⁹—decreases the cost of betrayal by facilitating the comparison of options. Incommensurability claims, which are just a denial that such comparison is made easily or routinely and in the absence of consideration of the intrinsic value of the options, emerge not only in the social and commercial games that we play with others, but in the psychological games that we play with ourselves.

⁴⁹ SIMMEL, *supra* note 1, at 128-29.