Berkeley Law

From the Selected Works of Robert Cooter

2008

The Misperception of Norms: The Psychology of Bias and the Economics of Equilibrium

Robert D. Cooter
Menachem Feldman, Hebrew University of Jerusalem
Yuval Feldman

Available at: https://works.bepress.com/robert_cooter/155/
The Misperception of Norms: The Psychology of Bias and the Economics of Equilibrium 1

ROBERT COOTER, MICHAL FELDMAN, YUVAL FELDMAN
UC Berkeley; Hebrew University of Jerusalem; Bar-Ilan University

This study combines the psychology of bias and the economics of equilibrium. We focus on two of the most discussed perceptual biases found by psychologists who studied the role social norms in ethical decision making. First, psychologists found a general tendency of people to over-estimate how many other people engage in unethical behavior. We show that this bias causes more people to violate the norm than if the bias were corrected. Second, psychologists found a general tendency of a person to over-estimate how many other people act the same as he does. We show that this bias does not change the number of people who violate the norm, contrary to the predictions of some psychologists. When a person suffers from both biases, they can augment or undermine each other. In either case, we show that supplying accurate information will cause more people to conform to the norm. In general, we show that applying the equilibrium concept to psychological studies improves public policy recommendations by increasing precision in predicting aggregate behavior over time.

1. INTRODUCTION

This study combines the psychology of bias and the economics of equilibrium. We will show that applying the equilibrium concept to psychological studies of cognitive biases increases precision in predicting aggregate behavior and reduces errors in making public policy recommendations.

Psychologists have found two systematic biases in the perception of norms in various contexts, including trade secrets, tax compliance, alcohol abuse, drugs, smoking, and water conservation. First, a person tends to over-estimate how many other people break social norms. We call this biased perception “Others are bad.” Second, a person tends to over-estimate how many other people act

1 Earlier versions of this paper were presented to the American Law and Economics Association, Harvard University (May 2007) and to a conference on Social Norms and Legal Compliance, May 2007, Israel. We would like to thank Dhammika Dharmapala, Adi Eyal, Rob MacCoun, Tom Tyler, and the participants of the above-mentioned conferences for their most useful comments.
like he does. We call this biased perception, “Others are just like me.” Here are two examples of these biases:

**Trade secrets:** Law prohibits an employee from disclosing the company’s trade secrets to others, although the law is vague about what counts as a trade secret. An employee in Silicon Valley resigns from her job and accepts a new job in a competing firm. She is inclined to share information with her new employer about her old employer. She believes that the inclinations of most employees in Silicon Valley resemble hers (“just-like-me”), and she believes that most other people break laws that she obeys (“others-are-bad”). Although she fears a suit, she decides that she will be safe if she behaves like her co-workers.

**Heavy drinking:** A college rule forbids drunkenness. A freshman in college wonders whether or not to try drinking alcohol heavily. He is inclined not to do it, and he believes that most other freshman share his inclination (“just-like-me”). He also believes that most students break social norms that he obeys (“others-are-bad”). He believes correctly that his social acceptance in the college depends on behaving like most other students.

In the psychological literature on perception of social norms, the belief that others are bad (or more accurately worse than they are in reality) is related to phenomena such as 'Uniqueness Bias,' Pluralistic Ignorance, the Norm of Self Interest and the Fundamental Attribution Error (see Goethals et al., 1991). The belief that others are just like me is mainly related to the 'False Consensus Effect' (Prentice and Miller, 1990). These phrases point towards psychological explanations of why we think that others are bad or just like us. Psychologists disagree about the explanation, so other phrases also appear in the literature. We avoid the psychological phrases in order to remain agnostic with respect to the causes. As explained, this paper concerns the effects of believing that others are bad or that others are just like me, not the causes.

---

2 It should be noted that there are a number of perspectives on the correct definition for those biases. Goethals and Klein (2000) seem to make a distinction between False Uniqueness and Uniqueness Bias, whereas Miller (1999) does not seem to make any distinction regarding how people underestimate the probability that other people will engage in ethical and morally desirable behavior (or call it "the tendency for people to underestimate the commonness of their own desirable behaviors (false uniqueness, as defined by Goethals 1986)," and overestimate the commonness of their own undesirable behaviors (uniqueness bias, as defined by Goethals, 1986). To test for this bias, participants are asked to estimate the mean of undesirable behavior and compare their estimates to the actual mean (for reviews, see Goethals, 1986; Goethals et al., 1991). See part 3 below for a detailed discussion of the relevance of these phenomena to the overestimation of unethical behavior in society.
We will use the preceding examples to explain in words the results of our mathematical model. In our two examples, the belief that others are bad makes the actor over-estimate how frequently others reveal trade secrets or drink heavily. The actor’s response to misperception depends on whether he sees an advantage or disadvantage from behaving like others. Where social norms are concerned, people often want to act like others, partly to gain approval and partly to reduce the likelihood of suffering a sanction for breaking the norm. If actors believe that safety lies in numbers, then the belief that others are bad will cause more actors to reveal trade secrets or drink heavily. In general the belief that others are bad causes more people to violate the norm than if the bias were corrected. Researchers who found others-are-bad bias with respect to protection of trade secrets, tax compliance, alcohol abuse, and water conservation, were right to predict that it would cause more people to do wrong.

Now we turn to the belief that others are just like me. This bias makes an actor who is inclined to reveal trade secrets or drink heavily overestimate how many other people do these things. Expressed normatively, actors who engage in “good” behavior overestimate the percentage of people in society who engage in good behavior. Conversely an actor who is inclined to refrain from these acts overestimates how many other people refrain. Expressed normatively, actors who engage in “bad” behavior overestimate the percentage of people in society who engage in bad behavior.3

Instead of changing behavior, just-like-me bias strengthens the inclinations that people already have. The person who is inclined to do wrong thinks that he has safety in numbers, which increases his resolve to do wrong. The person who is inclined to do right thinks that wrongdoers lack safety in numbers, which increases his resolve to do right. So just-like-me bias does not change the aggregate number of actors who do right and wrong. Researchers, who found just-like-me bias with respect to water conservation, smoking, and drugs, were wrong to predict that this bias will cause more people to do wrong.

While just-like-me bias does not change aggregate behavior, it increases its stability. For example, the bias causes fewer actors to change their behavior in response to an increase in the severity of a sanction for wrongdoing. Note, however, that just-like-me bias can also cause each wrongdoer to do wrong more often. In these circumstances, just-like-me bias increases the aggregate number of wrongs without increasing the aggregate number of wrongdoers.4

---

3 A famous exception is Prentice and Miller’s (1996) discussion of students’ overestimation of the number of students who engage in over-drinking.

4 We focus on the number of people who do right or wrong, which economists call the “extensive margin.” We could change the model to focus on how many wrongs are committed
What happens when a person has both biases? Do they augment or undermine each other? The answer is different depending on whether the actor is inclined to do wrong or right. For an actor inclined towards wrongdoing, just-like-me bias reduces his perceived risk of a sanction. Others-are-bad bias has the same effect. In this case, the two biases augment each other. Conversely, for a person inclined towards right doing, just-like-me bias increases his perceived risk of a sanction if he does wrong. Others-are-bad bias reduces his perceived risk of a sanction if he does wrong. In this case, one bias undermines the other.

We have explained that when a person suffers from both biases, they can augment or undermine each other. In either case, we show that supplying accurate information will cause more people to conform to the norm. Biased beliefs cause miscalculations that can impair a person’s performance. To reduce the risk of miscalculation and correct bias, a rational actor will sacrifice something of value – money, effort, or time – to obtain more information. As the cost of information approaches zero, bias tends to disappear among rational actors. Public policy can reduce bias by lowering the cost of information about the true frequency of wrongdoing. Given others-are-bad bias, whether on its own or in combination with just-like-me bias, public policy to disseminate the facts should cause a decrease in the number of wrongdoers. Thus public advertising of the facts about the level of wrongdoing should reduce the amount of it.

We have explained that others-are-bad causes more actors to do wrong. Our equilibrium analysis will show two distinct ways that others-are-bad produces this result. Others-are-bad can cause a few more people to do wrong, which is a marginal change. The economic analysis of wrongdoing primarily concerns marginal deterrence. Alternatively, others-are-bad can cause a large number of people to change their behavior. Most social norms have multiple equilibria, and a jump from one equilibrium to another causes a large change in the number of wrongdoers. The economic analysis of expressive law especially concerns jumps from one equilibrium to another. We will explain the technical conditions that determine whether a policy will produce a marginal change or a large jump. Marginal changes are more common and jumps between equilibria are more dramatic.

Part 2 documents research on the misperception of social norms. Part 3 describes its possible causes. Part 4 explains how to aggregate individual bias and predict the effects on equilibria. Part 5 discusses some policy conclusions, including mistaken conclusions by psychologists.

by each wrongdoer, which economists call the “intensive margin.” The bias we discuss, changes the former and not the latter.
2. EVIDENCE FOR THE MISPERCEPTION OF NORMS

We will describe psychological research that found misperception of norms. One of this article’s authors gave questionnaires to engineers in Silicon Valley concerning divulgence of trade secrets. The questionnaires asked each person whether or not he would violate trade secrets law, and then asked the frequency with which he thought that other people violated trade secrets laws. 44.8% of the study participants said that they were more likely than not to violate trade secrets law, but they estimated on average that 57% of the employees in their company would violate trade secrets law. When asked about the proportion of employees in Silicon Valley in general who would violate the trade secrets law, the average answer was 68%. Others-are-bad bias would produce such a gap in results.5

The gap between self-reported and perceived disclosures of trade secrets differed systematically across types of people. Those who reported that they were more likely to disclose secrets estimated that a relatively high percentage of other people disclose secrets, and those who reported that they were less likely to disclose secrets estimated that a relatively low percentage of other people disclose secrets. Just-like-me bias would produce such results.

The survey also found that the longer a worker spends in Silicon Valley, the more he feels justified in disclosing trade secret. Perhaps people change their beliefs to align with their misperception of the facts -- they accept the morality of the actual as they misperceive it.6

Research on tax evasion yielded similar results (see Wenzel, 2005). Psychologists have investigated the connection between the willingness of people to pay taxes and their perception of tax evasion by others.7 A longitudinal panel study of Australian taxpayers found that many individuals think that more people evade taxes than do in reality (Wenzel, 2005). Others-are-bad bias could produce this discrepancy.8

The same investigation found that the taxpayers' personal views of the morality of tax compliance affect their perception of the levels of tax compliance by others. Those with high personal standards of tax compliance

---

5 Interestingly, when the question has focused on the most severe type of trade-secret violation—intentional downloading of confidential know-how, only 20% said that they would download trade secrets and use it in a different company, but the participants believed that 43% of their peers would do so, and that 62% of the employees in Silicon Valley would do so.

6 For a discussion of the concept of 'conventional morality and the relationship between perceived norms and morality, see for example, Wood et al. (1988).

7 For a discussion of the concept of conditional cooperation, see Tyran and Feld (2002). For an illustration of the importance of conditional cooperation in the context of tax compliance, see Scholz and Lubell (1998).

8 Although in that study, the psychological cause they focused on was the pluralistic ignorance phenomenon.
perceived relatively more compliance by others, and those with low personal standards perceived relatively less tax compliance by others. These results are consistent with just-like-me bias.

Drinking, drugs, and smoking have been researched. A series of studies show that college students perceive their peers as less critical of heavy drinking than they actually are, which suggests others-are-bad. One of the studies tracked how attitudes developed over the course of two months among college freshmen and discovered gender differences. Male students adjusted their personal attitudes over time to match more closely the perceived consensus. After the adjustment, actual attitudes among males were closer to perceived attitudes. With female students, attitudes remained stable.\textsuperscript{9}

In a classical study on drugs, a sample of adolescents was divided into three groups: non-users, cannabis-users, and cannabis and amphetamine users (see Wolfson, 2000). The perceptions of members of the three groups differed significantly from each other.\textsuperscript{10} Compared to non-users, drug-users gave relatively high estimates of the number of users. These results are consistent with just-like-me bias.

Turning to smoking, researchers surveyed 916 junior high school students in Grade 7 and two years later in Grade 9 (Botvin et al., 1992). 7th graders who smoked thought that relatively more people smoke, and 7th graders who did not smoke thought that relatively fewer people smoked.\textsuperscript{11} Specifically, adolescents who were most involved with smoking believed that half or more than half of all adults or peers smoked, while those least involved believed that fewer than half of adults or peers smoked. Just-like-me bias could explain this data.

Now we turn to cooperation in a social dilemma situation. A tropical storm caused a water shortage in the eastern United States in 1999. Princeton University imposed a ban on showering for the first 3 days of the water crisis (see Monin and Norton, 2003). Princeton psychologists surveyed students for self-reported and

\textsuperscript{9} Subjects' ratings of personal comfort with Princeton drinking norms as well as estimates of their perceptions of the average students' comfort level were taken at two separate points separated by two months. The results indicated that men behaved in accordance with social influence theorist expectations in that they changed their own attitudes in the direction of the social norm. Indeed, personal average comfort among men in September was 5.84 with perceived average student at 7.48. In December, men's personal comfort with drinking norm jumped to 7.08, with perceptions of average student comfort level remaining relatively stable at 7.58. Women, on the other hand, showed no change in personal attitudes over time, save for a slight inflation of perceived average student comfort from 7.16 to 7.74.

\textsuperscript{10} Nonusers: M = 37.23\%, cannabis users: M = 44.38\%, and amphetamine users: M = 54.22\%.

\textsuperscript{11} The use of longitudinal and cross sectional design allows the authors to speak of causality, although they use correlation-based analysis.
perceived showering.\textsuperscript{12} During the ban, students estimated that others took more showers than implied by self-reports -- 47\% versus 33\% on day 2, and 56\% versus 47\% on day 3. Others-are-bad bias would have produced this gap.

As soon as the ban was lifted, other students were seen as taking fewer showers than implied by self-reports -- 70\% versus 77\% on day 4, 72\% versus 84\% on day 5. This fact is consistent with the theory that others-are-bad only applies to morally relevant behavior. During the ban, not showering was altruistic. After the ban was lifted, showering ceased to be morally relevant (or, perhaps showering became altruistic).

In another study on altruism, subjects were asked whether or not they would perform hypothetical acts to help others, such as aiding an aging couple stranded in a storm with a flat tire, and whether or not they thought that other people would perform those acts (see Goethals, 1986). The study found a gap suggesting a perceptual bias in the estimation of the proportion of wrong-doers in society. Others-are-bad bias was greatest for people who reported that they would not help others, which suggests just-like-me bias. Similarly, in the context of blood donations, Goethals (1986) found that 60\% of a student sample said they would be willing to donate blood but estimated that only 39\% of their peers would do so.

3. CAUSES OF PERCEPTUAL BIAS

We will review several accounts of the possible causes of others-are-bad bias, and then turn to the possible causes of the bias. The causes of this bias are both cognitive and motivational, and we review each type in turn. The media provides one of the simplest cognitive explanations for overestimating wrongdoing.\textsuperscript{13} Immoral acts get more media coverage than moral acts. Lichtenberg and MacLean (1992) demonstrate that much of what the media reports is bad news.\textsuperscript{14} Prentice and Miller analyzed four different types of TV news programs (national network news, local news, independent news, and cable network news) for violence, conflict, and suffering (see Johnson, 1996). Over

\textsuperscript{12} For the relationship between self-reported behavior and actual behavior, see Epley & Dunning (2000).

\textsuperscript{13} While psychology research on TV violence tends to focus on entertainment TV, the daily news is even more heavily watched and also contains extreme and realistic violent content, See Prentice and Miller (1996:808) for a discussion of the role of the media of this phenomenon.

\textsuperscript{14} Koren and Klein (1991) also show this in their comparison of news coverage of two scientific studies, one on bad news and one on good news. The studies compared by Koren and Klein both investigated the relationship between radiation exposure and cancer. The bad news study showed an increased risk of leukemia in white men working at the Oak Ridge National Laboratory. The good news study failed to show an increased risk of cancer in people residing near nuclear facilities. Koren and Klein found that subsequent newspaper coverage was far greater for the study showing increased risk.

DOI: 10.2202/1555-5879.1222
100 programs were analyzed over a six-month period. The authors discovered an emphasis on news that is bad and violent.\textsuperscript{15} The media also depicts immorality more than morality, including lying, adultery, robbery, and fraud.\textsuperscript{16}

In psychology, “salience” refers to the ease of recalling events that are vivid and recent. Psychologists have found that people tend to overestimate the probability of salient events and underestimate the probability of non-salient events (Tversky & Kahneman, 1973). The media gives salience to wrongdoing, which causes people to overestimate its frequency.

Why does the news media give disproportionate attention to bad news? The media competes for peoples’ attention. Human beings apparently pay more attention to bad news than good news (see Slaun, 1993). According to the theory of automatic vigilance, individuals who respond to undesirable social stimuli gain an evolutionary advantage (see Pratto and John, 1991). As a result, undesirable stimuli attract more attention than desirable social stimuli.\textsuperscript{17} This is especially true concerning undesirable behavior by others (see Skowronski & Carlston, 1989).\textsuperscript{18}

Another possible cause of others-are-bad bias is a deeper bias in the way people explain behavior. Many psychological studies support the conclusion that people attribute too much undesirable behavior to fixed traits of character in the actor, and too little to the social situation (Gilbert and Malone, 1995). Buckley, Harvey, and Beu (see Halbesleben et al., 2004) suggest that this

\textsuperscript{15} On parameters of time allocation and amount of featured news stories, more than half (53.4\%) of the news displayed violence, conflict, and suffering. Bad news was also given greater emphasis in that it featured earlier in the programs. While local news broadcasted the most bad news, all four program types were found to emphasize bad news.

\textsuperscript{16} Jones (1997) cites a USA Today report that 48\% of workers confess to illegal or unethical behavior as evidence of this phenomenon.

\textsuperscript{17} There is a fundamental asymmetry in people’s evaluations of gains and losses, of joy and pain, and of positive and negative events. A considerable body of research, in fields as diverse as decision making, impression formation, and emotional communication, has shown that people exhibit loss aversion (Kahneman and Tversky, 1984): They assign relatively more value, importance, and weight to events that have negative, rather than positive, implications for them. In decision making, potential costs are more influential than potential gains (e.g., Kahneman and Tversky, 1979). In impression formation, negative information is weighted more heavily than positive information (e.g., Fiske, 1980). In nonverbal communication, perceivers are more responsive to negatively toned messages than to positive ones (Fredi et al., 1978).

\textsuperscript{18} This concept is related to impression formation – the fact that unfavorable characteristics are weighted more heavily than favorable ones in initial formation of an assessment toward an object. Results showed that likableness ratings of a person associated with an unfavorable attribute were significantly more discrepant from a neutral impression than were ratings of a person associate with a desirable attribute. Additionally, subjects were more confident in their likableness ratings of the negative stimulus persons. It is certainly plausible to equate unethical traits/information with negative traits/information when it comes to impression formation.

Review of Law & Economics, © 2008 by bepress
“fundamental attribution error” plays a role in overestimating the amount of wrongdoing by others.\(^\text{19}\) According to the authors, most people are ethical with occasional lapses. Thoroughly unscrupulous people are an exceptional minority. The fundamental attribution error could cause a person who witnessed wrongdoing to conclude that the actor usually does wrong, whereas the correct conclusion in most cases is that the actor occasionally lapses.\(^\text{20}\)

Others-are-bad could also result from a tendency to believe that the behavior of others is instrumentally driven. The overestimation of unethical behavior could follow from a common belief that one's self interest is the most important factor in explaining the behavior of individuals in the society (Miller, 1999).

We have been discussing possible cognitive causes of others-are-bad bias. Another possible cause is mostly emotional. A wrongdoer may protect his self-esteem by exaggerating how frequently others commit the same wrong.\(^\text{21}\) Relevant concepts invoked by psychologists include social validation, self-enhancing biases, and constructive social comparison.\(^\text{22}\)

Now we turn from others-are-bad to just-like-me bias. An individual who projects his own behavior onto society overestimates how many others behave like he does.\(^\text{23}\) This bias is closely related to what the psychology literature calls the false consensus effect (FCE), which refers to a situation where people mistakenly think that others agree with them. According to the FCE, people tend to overestimate the social support of their own views and underestimate the social support for people who hold opposing views.\(^\text{24}\) Evidence from four studies in the original research by Ross demonstrates that social observers tend to form a false consensus with respect to the relative commonness of their own behavior. These results were obtained in questionnaires that presented subjects with hypothetical situations and also in actual conflicts that presented subjects with choices.

\(^{19}\) The fundamental attribution error is the basis of the pluralistic ignorance phenomenon, implying that people might fail to account for others true motivations for engaging in various behaviors.

\(^{20}\) For an economic model that develops the effect of attribution biases, see Dharmapala and McAdams (2005).

\(^{21}\) See Wenzel on taxation (2005).

\(^{22}\) Avoiding or inventing social reality when one suspects that the social practice might prevent him from following his own self-interest (Suls, 1987). See also Wheeler (1966).

\(^{23}\) For the false consensus effect, see generally Newman et al. (1997).

\(^{24}\) See Ross et al. (1977). But see also De la Haye (2000), demonstrating the mistakes that Ross et al. made in their measurements. In addition, Dawes and Milford (1996) also question the existence of the false consensus effect, basically showing the opposite, i.e. that the more people thought others were like them, the more accurate they were in their estimations of their behavior. See Marks (1984) for a reconciliation of the uniqueness and the false consensus effect. It is important to note that some have argued that the false consensus is not at all a bias, but could rather be explained according to the bayesian analysis.

DOI: 10.2202/1555-5879.1222
Several psychological mechanisms could cause just-like-me bias. One such mechanism is cognitive: A person may attend to positions with which he agrees and dismiss positions with which he disagrees. Selective attention allows his preferred position to dominate his consciousness. People also tend to associate with others who share their general beliefs, attitudes, and values. The number of arguments that we hear for or against something affects our attitudes towards it, and we hear more arguments from people inside our group than from outsiders, so the sorting of people reinforces selective attention. The association could be voluntary as when people select their friends, or involuntary as when people are segregated. If like associate with like, then recalling instances of similar behavior to your own will be easier than recalling dissimilar behavior.  

Instead of cognition, emotion could cause just-like-me bias. Most people presumably need to see their own acts, beliefs, and feelings as morally appropriate (Sherman et al., 1984). Finding similarity between oneself and others may validate the appropriateness of behavior, sustain self-esteem, restore cognitive balance, enhance perceived social support, or reduce anticipated social tensions.

4. THE MODEL

We will construct an economic model of conformity to a social norm, solve for the equilibrium, introduce perceptual bias into the model, and see how the equilibrium changes. We follow the economic tradition of distinguishing between benefits and costs. A person who breaks a norm often enjoys various benefits, such as the financial gain from disclosing trade secrets, the reduction in taxes from evading them, the pleasure of listening to music after downloading it illegally, victory from winning a contest by cheating, time saved by not complying with law, etc. Assume that each actor’s benefit from breaking the norm can be measured. The metric may be utility, pleasure, income, time, prestige, power, comfort, etc. In Figure 1, the vertical axis measures the amount a person benefits from breaking the norm. Each person i has a type θ, reflecting the benefits he obtains from breaking the norm. The horizontal axis depicts the cumulative proportion of people who enjoy a benefit of a given amount. According to the curve in Figure 1, a small number of people enjoy a high benefit, and a large number of people enjoy at least a small price.

We connect the benefit from breaking a social norm to standard economic concepts. A person’s benefit in economics is described as his “willingness to pay” (WTP). The curve in Figure 1 thus depicts willingness to pay for

---

26 See generally Marks and Miller (1987).

Review of Law & Economics, © 2008 by bepress
wrongdoing in a population of people. The number of people who are willing to pay a certain amount also measures demand. The curve in Figure 1 thus depicts the “demand” for wrongdoing. The demand curve slopes down because more people are willing to pay the price of wrongdoing as it decreases.

Now we turn from the individual’s benefits to his costs. Breaking a social norm often provokes social sanctions that can take various forms. First, people who break the norm could lose the social approval of their peers (McAdams, 1997). Second, they could face social resentment (Cooter, 1996). Third, they might have trouble finding business partners (Bernstein, 1992). Fourth, if the social norm is also a law, the wrongdoer might suffer civil liability or criminal punishment. Fifth, they might suffer in some or all of these ways because they acquire a bad reputation (Posner, 2000).

The vertical axis in Figure 2 indicates the individual actor’s cost of breaking the social norm, and the horizontal axis indicates the proportion of wrongdoers who break the norm. As depicted in Figure 2, the individual actor’s cost of breaking the norm decreases as the proportion of wrongdoers

![Figure 1: Distribution of the Willingness to Pay to Break the Norm](image1)
![Figure 2: Cost of Breaking the Norm as a Function of the Proportion of People Who Break the Norm](image2)

---

27 A number of factors moderate the relationship between violating a norm and social sanctioning (such as legitimacy of the law, morality etc. For a general discussion of people's motivation to sanction violators, see Falk et al. (2005).

DOI: 10.2202/1555-5879.1222
increases. Various reasons could explain why breaking a norm costs less when many others do it. A simple reason that is central to the enforcement of norms concerns the expected sanction. The expected sanction equals its probability multiplied by its severity. As discussed above, the probability that a particular wrongdoer will suffer a social or legal sanction often decreases as more people commit the sanctioned act. For example, when people see many others smoking in airports, they feel more confident that they will not be confronted if they smoke. Because safety lies in numbers, the cost curve slopes down in Figure 2. The curve in Figure 2 also describes the “supply of sanctions for wrongdoing” as a function of the proportion of wrongdoers.

Having characterized the demand and supply for wrongdoing, we can analyze the equilibrium towards which social interaction tends. We assume that the severity of the sanction suffered by wrongdoers can be measured by the same metric as the gain from wrongdoing, so the curves in Figure 1 and Figure 2 can be combined in the same graph. Consider a population of rational people with heterogeneous willingness to pay to break the norm, where person $i$’s WTP is denoted by $\theta_i$. Each person decides whether or not to break the norm by comparing the cost of the sanction and her WTP, $\theta_i$. We assume that the cost curve slopes down as in Figure 2. For example, suppose that the actor’s cost of violating the norm is inversely proportional to the fraction of people who break it. When a fraction $\lambda \in [0,1]$ of the agents violate the norm, the sanction cost is $1/\lambda$. Therefore, the net benefit $u_i$ of a person who breaks the norm is the difference between the benefit $\theta_i$ that he gains from his misbehavior and the cost $1/\lambda$ imposed on him, expressed by

$$u_i = \theta_i - 1/\lambda.$$ 

In contrast, an individual who complies with the norm obtains a utility of 0 (since he does not gain any benefits and does not incur any costs). Thus, the decision of a rational person with type $\theta_i$ is:

- Violate, if $\theta_i > 1/\lambda$
- Comply, otherwise.

Even within this minimal framework, we can derive some interesting implications. The percentage $\lambda$ of violators in equilibrium is determined as the intersection of the WTP distribution and the cost curve. In order to determine the equilibrium points, we need to make some assumptions on the WTP distribution in the population. For simplicity, we assume that an actor’s type is a random variable, distributed uniformly between 0 and some maximal value $\theta_m$. By this assumption, the cumulative proportion of actors who are willing to
pay a given price forms a straight line that slopes down, as depicted by the WTP curve in Figure 3.

Figure 3: The intersection points of the two curves represent two equilibria. The higher equilibrium, \( x_1 \), is stable, and the lower point, \( x_2 \), is unstable. The point \( x=0 \) is an additional stable equilibrium.

Figure 4: The intersection points of the two curves represent two equilibria. The higher equilibrium, \( x_1 \), is unstable, and the lower point, \( x_2 \), is stable. The point \( x=1 \) is an additional stable equilibrium.

Figure 3 depicts three equilibria. The two interior equilibria occur at the two intersections of the WTP curve and the cost curve, denoted \( x_1 \) and \( x_2 \). Thus at \( x_1 \) the cost of breaking the norm is too high for any of the \( f \cdot x_1 \) actors to break it, and the cost of breaking the norm is not high enough for any of the \( x_1 \) actors to stop breaking it. So the expected sanction at \( x_1 \) exactly sustains the proportion of wrongdoers and right-doers. The same is true for \( x_2 \). The third equilibrium is at \( x=0 \), where no one breaks the norm. When no one is breaking the norm, the expected sanction from breaking it is too high for anyone to change his behavior.

Consider the natural best-response dynamics of such a community. Starting at some initial \( x \) users arrive at individual decisions, their aggregate decisions define a new \( x \), which leads to a new aggregate decision, and so on. When the community is out of equilibrium, the direction in which the community moves depends on the relative heights of the WTP curve and the cost curve. If the cost curve lies above the WTP curve, violation cost is higher than current wrongdoers are willing to pay, so the fraction of violators decreases. This happens in Figure 3 for \( x<x_2 \) or \( x>x_1 \). In contrast, if the cost lies below the WTP curve, current wrongdoers are willing to pay more than the cost of wrongdoing, so the fraction of violators increases. This happens in Figure 3 for \( x_1<x<x_2 \).

According to these best-response dynamics, \( x=x_1 \) and \( x=0 \) are "stable" equilibria: If the population starts at a point close enough to these points, the
population will converge to these points. Conversely, \( x_2 \) is an "unstable"
equilibrium: If the population starts at a point around \( x_2 \), it will cause the
population to converge to one of the stable equilibria. If the initial \( x \) lies above
\( x_2 \), the process converges to \( x_1 \), and if the initial \( x \) lies below \( x_2 \), it will cause
the population to converge to \( x=0 \).

The population is unlikely to persist at the unstable equilibrium \( x_2 \), so we will not
focus on it. In addition, since few social norms enjoy compliance approaching
100\%, we will not focus on \( x=0 \) either. Therefore, we restrict attention to the
stable interior equilibrium, \( x_1 \), and refer to it as the most likely equilibrium in the
population. In addition, we will refer to the second point, \( x_2 \), as the \textit{tipping point}
above which the population will converge to the equilibrium \( x_2 \).

The properties of the equilibria described above depend on the shape of the cost
curve and the distribution curve, especially whether the cost curve intersects the
WTP curve from above or below. For example, if the cost and WTP curves
intersect in the opposite direction of Figure 3, as in Figure 4, the stable equilibria
are the interior one, \( x=x_2 \), and the corner one, \( x=I \), in which everyone violates
the norm. In Figure 4, either most actors obey the norm or else no one obeys it.

In standard economic analysis of markets, the demand curve slopes down as in
Figure 1. In standard economic analysis, however, the supply curve does \textit{not} slope
down as in Figure 2. The fact that both curves slope down creates the possibility
of multiple equilibria as depicted in Figures 3 and 4. The supply curve slopes
down for social norms when an increase in the number of wrongdoers decreases
the cost that each one of them expects to bear. We believe that this is the
common situation. Consequently, we believe that multiple equilibria are more
common for social norms than in markets for commodities. Given multiple
equilibria, small changes in the social environment can produce abrupt jumps up
or down in levels of compliance with norms. For example, in Figure 4, a small
change from \( x_1 \) will cause the system to jump to \( x_2 \).

\footnote{An implication of the existence of multiple equilibria is that policymaking requires global
knowledge of the cost function and the distribution of the willingness to pay. This is due to the
fact that a minor movement may initiate a jump in the aggregate behavior. This is in contrast to
standard economics, where there exists a unique equilibrium, and thus policymaking requires
only a local knowledge of the demand and supply curves.}
While we believe that a down-sloping cost curve is more common for social norms, cases also occur where the cost curve slopes up. Instead of gaining an advantage from acting like others, sometimes one gains an advantage from acting differently from others. Instead of safety in numbers, there is sometimes danger in numbers. This possibility has many causes. Thus, as wrongdoing increases, society may shift so many resources into punishing wrongdoers that the probability increases that a wrongdoer will suffer a sanction. Consider the case of drunk driving. It is a more pervasive problem in the U.S. than in Israel, so police in the U.S. buy state-of-the-art detection equipment that Israel foregoes. Better equipment may make the likelihood of detection higher in the U.S. than in Israel.

The fixpoint dynamics in this case where the cost curve slopes up are depicted in Figure 5. If $x<x_1$, then the cost is lower than the WTP, so the number of violators increases. If, however, $x_1<x$, then the cost exceeds the WTP, so the number of violators decreases. According to these dynamics, $x=x_1$ is the unique equilibrium, and it is a stable one. When the cost curve slopes up and the demand curve slope down, small changes in the social environment cause gradual changes in the levels of compliance.

4.1. OTHERS-ARE-BAD BIAS
Now we use our model to analyze the two biases that we already described. The others-are-bad bias suggests that the individual's perception about the fraction of people who break the norm is biased upwards. When individuals overestimate the percentage of people who break the norm, the cost curve
shifts down, which Figure 6 depicts as the shift from the “real cost” curve to the “biased cost” curve. A downward shift in the cost curve causes a higher fraction of people to break the norm in equilibrium, as indicated by the shift in the stable interior equilibrium from $x_1$ to $x_{1b}$. Thus others-are-bad causes a larger proportion of people to violate the norm.

Now consider the dynamics of convergence to the equilibrium. Specifically, consider the circumstances under which society converges to the corner equilibrium where no one breaks the norm. With others-are-bad bias, we need a higher fraction of people who initially comply with the norm in order to arrive at the equilibrium where no one breaks it. If the initial $x$ lies below $x_{2b}$ and there is others-are-bad bias, the community will converge to the equilibrium $x=0$ where no one breaks the norm. For initial values of $x$ in between $x_2$ and $x_{2b}$, unbiased perception causes the community to settle where no one breaks the social norm, whereas biased perception causes the community to settle at $x_{1b}$ where most people break the norm.

In summary, others-are-bad bias has two effects on the equilibria:

- The equilibrium where no one violates the norm is less likely.
- The likely equilibrium is one where more people violate the norm.

Now we turn to bias reduction. Others-are-bad bias causes some actors to break the norm who objectively stood to gain more from conforming to it. A rational actor recognizes that bias can lead to costly mistakes, so he will pay to reduce his own bias. Consider the possibility that actors can obtain information to reduce or eliminate bias. We will formulate how much a rational actor would pay to reduce his possible bias.

An actor’s net benefit from conforming to a norm is normalized to equal zero. An actor’s benefit from breaking the norm is $\theta_i$, and the actual cost of breaking it is $1/x$, so his actual net payoff from breaking the norm is $\theta_i-1/x$. Let $x_{i'}$ denote person $i$’s perception of the fraction of wrongdoers. The actor’s perceived net payoff from wrongdoing is $\theta_i-1/x_{i'}$. Because of the others-are-bad bias, the perceived fraction of wrongdoers is higher than the actual fraction, $x_{i'}/x > x$, so the perceived net payoff from wrongdoing is higher than the actual net payoff. Thus others-are-bad bias causes some actors to break the norm who stand to lose from doing so.\footnote{By definition, the marginal actor’s benefit from breaking the norm equals his perceived cost: $\theta_i = 1/x_{i'}$. Consequently, others-are-bad bias causes the marginal actor to break the norm even though his actual net gain from doing so is negative.} Information that caused such an actor to switch from wrongdoing to right doing would save him $1/x - \theta_i$. Suppose that an actor knows that his perception might be biased. Let $q$ denote the actor’s subjective probability that receiving the
Psychology of Bias and the Economics of Equilibrium / 905

information will cause him to switch from wrongdoing to right doing. He will be willing to pay for the information only if its cost, denoted \( a_i \), is less than his expected savings from switching behavior,\(^{30}\) that is, if:

\[
q \left( \frac{1}{x} - \theta_t \right) \geq a.
\]

Let \( \hat{a} \) denote the most that anyone would pay for information.\(^{31}\) If the information cost is higher than \( \hat{a} \), then no one buys information and the equilibrium \( x_t^P \) based on biased perception will persist. As the price of information falls below \( \hat{a} \), more and more actors buy and correct their bias. As the price of information falls to \( \theta \) for a population of rational actors, everyone will buy it, their bias will disappear, and society will approach the equilibrium \( x_t \) based on accurate perception. Figure 7 depicts the situation for a population of rational actors. If the cost of information is in the range between \( \theta \) and \( \hat{a} \), the equilibrium will lie somewhere between \( x_t \) and \( x_t^P \).

Figure 7: This figure demonstrates the effect of the information cost on the interior stable equilibrium. If the information cost is 0, then all the agents are informed about the true distribution, and therefore the original equilibrium \( x_1 \) is the only equilibrium. If, however, information cost is high, agents will not buy it, and the worse equilibrium \( x_1^P \) will persist.

\(^{30}\) For example, assume that others-are-bad causes the perceived cost of norm-breaking to equal \( 1/2x \), whereas the true cost is \( 1/x \). For the marginal actor, his WTP equals his perceived cost of breaking the norm: \( \theta_t = 1/(2x) \). The actual net payoff of the marginal actor from breaking the norm is \( 1/(2x) - 1/x = -1/(2x) < 0 \). If he did not break the norm, his actual net payoff would be 0. So he would gain from spending up to \( -1/(2x) \) to obtain information about the true cost of breaking the norm.

\(^{31}\) Under reasonable assumptions, the marginal wrongdoer will pay the most for information. The marginal wrongdoer has the lowest value benefit \( \theta_t \) among wrongdoers. Consequently, if \( q \)

DOI: 10.2202/1555-5879.1222
4.2. Just-Like-Me Bias

Now we turn from the others-are-bad bias to the just-like-me bias. To establish a benchmark, consider the full information, unbiased population that converges to an equilibrium in which a fraction \( x \) breaks the norm. If full information causes an actor to break the norm, we say that he is inclined to be a wrongdoer. For such a wrongdoer, the benefit \( \theta_i \) equals or exceeds his actual cost \( 1/x \) of breaking the norm: \( \theta_i \geq 1/x \). Now introduce just-like-me bias and see this value's change. Let \( x^p_i \) denote person \( i \)'s perception of the fraction of wrongdoers. By assumption person \( i \) is a wrongdoer, so the just-like-me bias causes \( i \) to over-estimate how many other people are wrongdoers. Thus we have

\[
\theta_i \geq 1/x \geq 1/x^p_i
\]

Bias thus increases the wrongdoer's perceived net benefit from violating the norm, so the wrongdoer does not change his behavior, although bias gives him a stronger reason to do what he does.

If, on the other hand, full information causes an actor to obey the norm, we say that he is inclined to be a right doer. For such a right doer, his benefit \( \theta_i \) is less than his actual cost \( 1/x \) of breaking the norm: \( \theta_i \leq 1/x \). We introduce the just-like-me bias, which causes \( i \) to under-estimate how many other people act the same as he does: \( x^p_i < x \). We have

\[
\theta_i < 1/x < 1/x^p_i
\]

Bias thus increases the right doer's perceived net benefit from obeying the norm, so the right doer does not change his behavior, although bias gives him a stronger reason to do what he does.\(^{32}\)

4.3. Combining the Two Biases

What happens when each actor is inclined to both biases? The answer is different depending on whether the actor would do wrong if he knew its actual cost. For such wrongdoers, both others-are-bad bias and just-like-me bias increase the perceived gain from wrongdoing, so they reinforce each other and wrongdoing persist. For right doers, the others-are-bad bias increased the

---

\(^{32}\) We do not discuss just-like-me bias in a person who is poised between right-doing and wrongdoing. Presumably such a person arbitrarily chooses whether to do right or wrong, and just-like-me bias makes him assume that others make the same choice, thus increasing his commitment to a choice that was originally arbitrary.

---
perceived net gain from wrongdoing, and the just-like-me bias decreases it, so they undermine each other. If the others-are-bad bias is larger, then the actor believes that the proportion of wrongdoers is larger than it really is, which is a reason to switch and do wrong. If the just-like-me bias is larger, then the actor believes that the proportion of wrongdoers is smaller than it really is, which is a reason to persist in doing right. Depending on the larger effect, the actor may persist in right doing or switch to wrongdoing. Thus when both biases combine, wrongdoers do not change their behavior and some right doers may switch to wrongdoing, so the aggregate level of wrongdoing may increase and will not decrease. Any increase in the aggregate number of people who do wrong is due to the others-are-bad bias, not to the just-like-me bias.

5. POLICY IMPLICATIONS
We combined the psychology of bias and our equilibrium model to deduce policy conclusions. The research found evidence of others-are-bad with respect to various contexts, such as disclosing trade secrets, evading taxes, drinking heavily, using drugs and smoking, and showering during a water emergency. The others-are-bad bias makes the actor's perceived cost curve for wrong doing lie below the actual cost curve in Figure 5. Consequently, our model predicts more disclosing of trade secrets, evading taxes, heavy drinking, drug use, smoking, and showering than would occur if actors' others-are-bad bias were corrected. Providing accurate information on the amount of wrongdoing should reduce the bias and diminish wrongdoing. Increasing recall of concrete examples of right doing or social sanctions against wrongdoing might have a similar effect.

We noted that the study of college drinking found that male students increased their reported drinking over time to match more closely the perceived consensus. In terms of Figure 5, the actual cost curve shifted closer to the perceived cost curve with time. At the semester's beginning there is more bias to correct, so corrective measures should have more effect. Thus accurate information about heavy drinking should be given to males when they first arrive at college. With female students, attitudes remained stable, so the timing of the information should make less difference to their behavior.

In the study of tax compliance, researchers were able to monitor people's actual tax files. Some sub-groups were given information about the gap between their own behavior and the behavior of others. Receiving information on the behavior of others caused more tax compliance in some forms, such as claiming deductions. This fact is consistent with our prediction that disseminating accurate information will cause more right doing when actors

DOI: 10.2202/1555-5879.1222
suffer from others-are-bad bias, regardless of whether or not they also suffer from just-like-me bias.

The study of showering showed that compliance with the ban decreased with time. These facts suggest a movement towards equilibrium at a point like \( x_1 \) in Figure 3 where most people violate the norm. Our model predicts that accurate information about wrongdoing would reduce the amount of wrongdoing in the final equilibrium.

Reduction in others-are-bad should usually cause a small change in aggregate behavior, which the tax study detected. Occasionally, however, correcting bias may cause a large jump in aggregate right doing. Thus an initial proportion of wrong doers below \( x_2 \) in Figure 3 will cause their numbers to fall approximately to zero, whereas an initial proportion of wrong doers above \( x_2 \) will cause their numbers to rise to \( x_1 \). In the study of alcohol, perhaps more male students drink heavily when they arrive as freshmen, which causes the system to settle at a high level of drinking among males. Conversely, perhaps fewer female students drink heavily when they arrive as freshman, which causes the system to settle at a low level of drinking among females. The researchers did not systematically explore the possibility of multiple equilibria, so we do not know whether or not they actually exist in this case.

According to our model, the just-like-me bias will not affect the number of wrong doers. Thus we predict that just-like-me bias will not cause more citizens to avoid paying taxes, more employees to disclose trade secrets, more freshmen to drink heavily, more secondary students to use drugs or smoke, or more college students to evade a ban on showering during a water emergency. Providing accurate information to correct just-like-me bias will not reduce the number of wrong doers, although information that corrects others-are-bad will have this effect.

While the just-like-me bias does not cause more wrong doers, it may cause more wrongdoing. Thus just-like-me bias should increase the total number of acts (but not the number of actors) of disclosure of trade secrets, tax evasion, heavy drinking by college freshmen, drug use and smoking by secondary students, and showering during a water emergency.

Just-like-me bias should also make people more reluctant to change their behavior. For example, just-like-me bias should blunt the effectiveness of punishment to deter wrongdoing. When just-like-me is more emotional than cognitive, bias resists correction by fresh information. Information is probably not enough to correct bias whose basis is more emotional.

The study of smoking found that students in 9th grade were more likely to smoke than in 7th grade. The study also showed which 7th grade non-smokers were likely to become 9th grade smokers. Specifically, non-smokers in the 7th
grade who thought that others smoke were more likely to become smokers in the 9th grade. In other words, the non-smokers who failed to project their good behavior onto others were more likely to become smokers. In contrast, 7th graders who thought that others do not smoke were relatively unlikely to smoke themselves in the 9th grade. Thus non-smokers who projected their good behavior onto others were unlikely to become smokers.

6. CONCLUSION

When psychologists find perceptual bias in individuals, they need analytical tools in order to predict the effects on aggregate behavior. We combine the psychology of bias and the economics of equilibrium to predict how psychological biases affect aggregate behavior. We analyzed two fundamental biases relevant to law and norms -- the belief that others are bad, and the belief that others are just like me. We focus on these two biases to the exclusion of others for two reasons. First, the empirical research on norms emphasizes their importance and pervasiveness. Second, a careful analysis of their effects is difficult and powerful. By working out the effects of two biases in detail, we hope that we convinced others that the power justifies the difficulty of applying the equilibrium concept to individual psychology.

References


Psychology of Bias and the Economics of Equilibrium / 911


DOI: 10.2202/1555-5879.1222