12-2005

Blind Spots: The Effects of Information and Stakes on Fairness Bias and Dispersion

James Konow
Loyola Marymount University, jkonow@lmu.edu

Recommended Citation

Konow, James (2005). "Blind Spots: The Effects of Information and Stakes on Fairness Bias and Dispersion," Social Justice Research, vol. 18, no. 4 (December), pp. 349-390.

April 2004

Blind Spots: The Effects of Information and Stakes on Justice Biases

James Konow
Department of Economics
Loyola Marymount University
One LMU Drive, Suite 4200
Los Angeles, CA 90045-2659
Telephone: (310) 338-2956
Fax: (310) 338-1950
Email: jkonow@lmu.edu

Abstract

Justice is commonly regarded as a prominent force in collective choice. Yet mounting empirical evidence demonstrates that beliefs about fairness exhibit an egocentric, or self-serving, bias that often impedes agreement among parties whose interests conflict. This study undertakes a comparative analysis using experimental evidence of the effects on fairness biases of personal stakes and of the quantity and quality of information. It finds a well-behaved trade-off between self-interest and fairness and, contrary to conventional wisdom, a favorable effect of information on impartiality and justice. Inhibiting personal stakes and increasing information reduces bias and increases convergence of fair beliefs and outcomes.

Keywords: Justice; fairness; equity; self-interest; bias
JEL classification: A13; D61; D63

I wish to thank the Ludwig Boltzmann Institute for financial support and Wulf Gaertner, Reinhard Neck, Joe Oppenheimer, Lars Schwettmann and participants of the meetings of the Public Choice Society for helpful comments and suggestions on this paper.

Though it may be true, therefore, that every individual, in his own breast, naturally prefers himself to all mankind, yet he dares not look mankind in the face, and avow that he acts according to this principle. He feels that in this preference they can never go along with him, and that how natural soever it may be to him, it must always appear excessive and extravagant to them.

– Adam Smith, The Theory of Moral Sentiments, pg. 126, 1759 (1809 edition)

1. Introduction

Along with self-interest, justice is commonly regarded as a prominent force in collective choice. Buchanan (1985) writes "It is important that (the) laws and institutions (of society) be seen to be fair… But, perhaps even more important, the institutions of political decision-making must also be seen to be fair and just" (pg. 138). Other writers echo these themes. Kirchgässner (2004) argues the necessity of just, or moral, behavior for the proper functioning of markets and the democratic political order. For Rawls (1971), justice is "the first virtue of social institutions" that informs their basic structure. Kolm (1996) argues for the full integration of justice into public choice and outlines how just, or moral, behavior can be induced with only a modicum of concern for it: if purely self-interested agents care about the approval of others and if others may express their moral approval or disapproval at negligible cost, self-interested agents will incur costs for morality. Indeed, in the public good experiment of van Dijk, Sonnemans and van Winden (2002), anonymous subjects demonstrate a willingness to incur greater than negligible costs to solve free rider problems, even without explicit social approval, and this impacts the social ties that can reinforce and be reinforced by social norms like justice. Nevertheless, as Buchanan points out, "questions of 'justice' necessarily get mixed up and intermingled with pure self-interest" (pg. 125). Various studies show that stakeholders suffer from "blind spots" about justice, i.e., they form beliefs about fairness that exhibit an egocentric (or self-serving) bias. When opposing parties act on such beliefs, it is not clear that the "right" amount of justice, or, indeed, anything approaching justice, will result. The current study analyzes evidence about egocentric fairness biases and about how they are affected by personal stakes and by information about the contexts in which they emerge.

It seems that justice is inseparable from self-interest in the literature. For example,

Rawls's social contract approach is premised on individuals who pursue their own interests in selecting the fundamental structure of society from an "original position" of ignorance about their ultimate places in that society. Buchanan and other scholars of the constitutional contract also adopt a contractarian approach involving self-interested agents, but, as Rowley (1993) points out, they avoid the strong and unrealistic informational constraints of Rawls. As the above quote suggests, Smith expects no magnanimity of people and, in fact, he heralds the virtues of self-interest in *The Wealth of Nations* (1776). Yet, he indicates that this motive cannot serve as a moral principle and admonishes readers to adopt the stance of impartial spectators in *The Theory of Moral Sentiments* (1759). Smith's impartiality construct suggests a means to reconcile these ostensibly dissonant objectives as well as a motivation for a study such as this one. Legitimate justice claims must satisfy impartiality, i.e., they cannot be biased by personal stakes. Combining this concern with the public choice approach to information, which treats individuals as at least partially informed, the current study explores how personal stakes and the quantity and quality of information provided to stakeholders affect two phenomena: the egocentric *bias* exhibited by parties and *convergence*, i.e., the degree to which parties agree on what is fair or are willing to act on a common view of fairness. Both bias and convergence are important to normative studies of justice, e.g., Rawls's construct rests on a particular concept of impartiality but also on the presumed consensus of people about what is just. These two variables also go straight to issues of concern for the analysis of constitutional contracts and of the use or misuse of justice in ongoing political and social behavior. Legitimacy in these areas is bolstered by impartiality and consensus. Moreover, actual agreement on rules of allocation increases, as we will see, with convergence of views about justice.

Can justice, though, impartially guide the rules or decisions of a society given the acknowledged impact of egocentric biases on social behavior? At first glance, evidence from the growing social science literature on the topic does not seem reassuring: empirical studies demonstrate a well-established pattern of self-serving biases when people are implicated parties. For example, 88% of Americans believe they are above average drivers (Svenson, 1981). Even

the educated succumb to this bias: 94% of college professors in one survey said they perform above average work (Cross, 1977). Egocentric biases about justice are implicated in a wide range of problems, including armed conflicts, labor strikes and legislative logjams. They present of source of variation in the data that can confound inferences by social scientists about impartial justice. Attempts to rectify injustice presume agreement on what justice demands, and egocentric fairness biases often prove a formidable obstacle. Numerous studies have identified these biases: when people are stakeholders, they tend to view their fair share of a reward as being greater than their fair share in the view of others (e.g., Messick and Sentis, 1983), and they reveal a willingness to act on biased views of fairness, including in experiments with monetary rewards (e.g., Babcock, et al., 1995, Konow, 2000) and in the field as real employers and employees (Babcock, Wang and Loewenstein, 1996). Knowledge about factors relevant to justice, on the other hand, provides the means to achieve justice more reliably, whether this information is for use by individuals who seek to promote justice or for consideration by adversaries in resolving their conflicting interests. An important open question is whether, on balance, increased information helps agents to achieve fairer outcomes or whether biased processing and use of such information contributes to less fair outcomes. These interrelationships between fairness bias, stakes and information have not been comprehensively investigated. This paper undertakes such a comparative analysis and reports a clear pattern of effects that can inform the analysis of public choice and offer a means to temper fairness biases.

The more detailed analysis in this paper centers on evidence from nine previously reported experiments involving real monetary stakes in which the subjects' interests conflict. These studies were selected because, for our purposes, they control stakes and/or information in a more focused manner than other experiments or than is possible in field studies. Data from these experiments are reinterpreted in a common format and undergo an original analysis for bias and convergence. The results suggest a well-behaved trade-off between self-interest and fairness. Nevertheless, parties often employ fairness concepts opportunistically, using the one that serves their interests. Dispersion of views and disagreements, therefore, tend to increase with conflicts

in interests and with the availability of competing concepts. Overall, however, providing parties with information relevant to justice tends to bring about outcomes that are less biased and less disperse. Studies that manipulate stakes or the timing of information about stakes indicate that the egocentric bias can be traced largely to self-deception, i.e., egocentric processing of information to bolster beliefs that it is fair to be unfair to others. Collectively, these studies suggest that views of justice lose their bias when people are disassociated from their interests while becoming informed of the relevant specifics, along the lines of Adam Smith's impartial spectator. This suggests a means for identifying unbiased justice while supplying a foundation for the descriptive and prescriptive analysis of individual and collective choice.

Section 2 reviews the foundational literature from survey studies of egocentric fairness biases, and section 3 examines the effects on fairness biases when stakes vary. The effects of information on biases are analyzed in section 4, and section 5 concludes.

2. Background

Numerous survey studies suggest that views of justice are related to the personal interests, experiences and characteristics of individuals. Many of these employ vignettes, or hypothetical scenarios, and elicit the preferences of respondents over alternate allocations. Gaertner, Jungeilges and Neck (2001), for example, investigated the preferences of students in several countries for the distribution of economic resources. In this survey, respondents make pair-wise choices between distributions (x, y) that differ with respect to the degree they satisfy basic needs, address a utilitarian concern for the welfare of largest number, encourage economic growth or infringe on human rights. The authors relate many of the patterns of cross-country differences in judgments to the circumstances of individuals in those nations. For instance, students in several economically depressed formerly socialist countries exhibited a greater concern for economic growth and promoting more advantaged persons than for helping the needy or protecting human rights. Alluding to the Rawlsian impartiality construct, they conclude that it is "obvious that people do not put themselves under a 'veil of ignorance' but consciously take into account their personal interests when making judgments" (pg. 963).

The above study drew on student respondents and asked them which distribution "should" be chosen. Another set of studies examined non-student populations and used explicit fairness terminology. Kahneman, Knetsch and Thaler (1986) presented randomly sampled telephone respondents in Toronto and Vancouver with hypothetical scenarios and asked them to evaluate whether or not the transaction terms in each of those questions were fair. Gorman and Kehr (1992) posed most of these same questions (22, counting contrasting versions of some scenarios) to a sample of business executives. Although the majority response (acceptable or unfair) in the two samples was the same for most of the questions, the percentage of respondents judging transaction terms unfair differed significantly across the two groups of respondents for 82% of the questions. In every one of those cases, this difference was associated with the business executives favoring the interests of the firm over those of the consumer or worker, relative to the more general population.

The study of Schokkaert and Capeau (1991) provides more thorough evidence on subject attributes that permits interpretations of personal bias. They elicited views of a representative sample of the adult Belgian working population on the fair compensation to workers in several hypothetical scenarios and collected information on various socio-economic characteristics of the respondents. Also, their survey contains five to seven quantitative response categories per question representing progressive differences in distributions as opposed to the more limited qualitative response formats of the Gaertner, Jungeilges and Neck (distribution *x* or *y*) and Gorman and Kehr (acceptable or unfair) studies. Certain patterns emerge, such as the tendency of blue-collar workers to reward direct contributions and of older people to defend seniority rights. Schokkaert and Capeau conclude that "There are clear indications that the self-interest of the respondents influences their answers" (pg. 342), although they find that this effect is not overwhelming.

The previously cited studies examine distributive preferences of subjects over the hypothetical allocations of others and permit one to associate these with personal characteristics of respondents. Those results suggest that subjects respond at least somewhat vicariously,

identifying with one party or another. Messick and Sentis (1983), on the other hand, directly made subjects hypothetical stakeholders and examined their biases. Specifically, student subjects were instructed to "imagine" that they worked a different number of hours (7 or 10) grading exams for a professor. The fairness wording was explicit, and the response format continuous. These studies reveal significant egocentric biases in the pay deemed fair. In particular, the use of fairness rules is self-serving: subjects presumed to work 7 hours were more than twice as likely to employ equal pay than those working 10 hours. On the other hand, those who worked more hours were more likely to apply proportionality of pay to hours worked.

In the above studies, respondents were asked to judge hypothetical scenarios with hypothetical stakes. An important issue is whether these results generalize to real stakeholders in the real world. As Messick and Sentis ask, "Could it be that the egocentric fairness bias is a phenomenon that is manifested only when people are filling out questionnaires about hypothetical situations and hypothetical payoffs?" (pg. 76). The study of Babcock, Wang and Loewenstein (1996) helps shed light on this matter as well as on the question of whether even experienced negotiators fall victim to the egocentric bias. They surveyed union presidents and school board presidents in Pennsylvania regarding what school districts they would consider comparable for purposes of teacher salary negotiations. The average salary in districts chosen by the unions was significantly greater than the average in districts chosen by school boards. This difference was also positively correlated with the variance in teacher salaries in neighboring districts, implying larger variation in comparison values provides more opportunity for egocentric evaluations. Moreover, districts in which differences in the two parties' estimates were greater were much more likely to have experienced strikes in the past, which suggests the bias leads to disagreements with real social and economic costs.

The studies analyzed in following sections were selected because they satisfy certain criteria that are useful given the goals of this paper. First, they all involve *salient rewards*, i.e., the stakes are not hypothetical but rather involve real monetary rewards, and individual payoffs depend in an important way on subject decisions. The work of Babcock, Wang and Loewenstein

helps address the concern of Messick and Sentis about hypothetical situations, but the decisions of the participants in that study did not affect any real payoffs, in contrast to the studies selected here. Although studies with and without real stakes often lead to similar conclusions, some results are affected by the presence of real stakes, such as the average level of unfairness. The monetary stakes also permit an unambiguous, interpersonally observable measure of different motives. Second, the stakes in each of the studies are fixed, which helps to focus the attention on equity issues, as opposed to preferences over the size of surplus. Third, the stakes are allocated between only two parties. This limits the role of more complicated strategic considerations that might confound inferences about justice. Fourth, the formats for decisions are continuous, or nearly continuous. Thus, allocations can be clearly expressed and interpreted on the unit interval, which permits comparative analysis of bias and convergence across studies. In this paper, bias is measured by differences in mean allocations. Convergence is evaluated by differences in standard deviations.

3. Stakes and Fairness

This section explores questions of how the willingness of agents to act on fairness varies with stakes. Does it matter whether the stakes are real or hypothetical? What effect does the size of stakes have? How do variations in the dictates of self-interest affect claims on stakes?

3.1 *Effects of Variation in Stakes on Claims*

Forsythe, Horowitz, Savin and Sefton (1994) conducted versions of a simple experiment called the *dictator game*. In the dictator game, each subject in one group, called a dictator, receives a fixed sum of money. The dictator may give any amount, or nothing, to an anonymous counterpart, whereby the counterpart otherwise receives nothing (except for a small show up fee that all subjects receive). Forsythe, et al. conducted three treatments of this game in a between subjects design (i.e., each treatment used a different group of subjects): 1. Hypothetical $5 Dictator, 2. Actual $5 Dictator and 3. Actual $10 Dictator. In the first treatment, dictators were asked how much of $5 they wished to share with their counterpart, but no one was actually paid the chosen amounts. The second treatment was similar, but dictator decisions resulted in actual

payments and subjects were told that. The third treatment was the same as the second, except dictators were given $10 to divide.

The results are summarized in the first rubric of Table 1. Results reported in this paper are normalized: the means and standard deviations are expressed as a fraction of the total stakes. These values refer to claims by the party that is typically advantaged in the sense of being the party that, in most treatments of the experiment, is able to claim more than one-half. Here, of course, that is the dictator, who is the *first mover* (indeed, in this experiment, the only mover). We see that, when the stakes are not salient and subject decisions do not result in real payments, dictators indicate on average that they will take 61.7% of the (hypothetical) $5. The superscript 0 on the mean of .617 indicates that this mean differs from equal splits of .5 at the 5 percent level of significance. Later in the paper we will address studies in which fair divisions differ from equal splits. I have argued elsewhere (e.g., Konow, 2003) that, although equality in some sense is related to justice, equality in outcomes is not a general principle of justice. Nevertheless, evidence from numerous studies show that in experiments, such as this one, where the context does not provide any informational basis for distinguishing claims of parties, subjects consider equal splits fair. Thus, the result from treatment 1 shows that subjects yield, to some degree, to self-interest and are willing to acknowledge their unfairness, even when there is no material benefit to doing so. Nevertheless, hypothetical dictator decisions underestimate the degree of unfairness exhibited by real stakeholders, as the second treatment demonstrates. When the $5 payments are real, dictators choose to keep 77.8% of the stakes. This differs significantly from .5, once again denoted by the superscript 0 on the mean, and from the mean in treatment 1, as indicated by the superscript 1 on the mean of .778 (as well as by the superscript 2 on the mean of .617 in the first treatment). Raising the stakes from $5 to $10, however, has no significant effect on dictator claims, given the mean of .767 in the third treatment. This mean differs significantly from .5 and from the mean in the first treatment, but not from the mean in the second treatment. In addition, all of the means reported in this and all other tables in the paper differ significantly from the claim of 1 that corresponds to purely self-interested behavior. These results suggest

that, on average, people trade off self-interest with fairness, but do not choose completely self-interested or completely fair allocations. In addition, hypothetical dictator decisions underestimate the magnitude of unfairness, but $5 and $10 stakes produce similar claims.

The next column in Table 1 lists the standard deviations of the claims. These do not differ significantly based on the stakes, suggesting that convergence is not affected by the magnitude of stakes involved. The final two columns show the percentage of claims that can be categorized as equal (.5) or self-interested (1). Using real versus hypothetical stakes causes the percentage of equal, or fair, claims to drop significantly from 48% to around 20% in treatments 2 and 3. Using real $5 stakes causes the percentage of completely self-interested claims to rise significantly from 13% to 36%. The increase in completely self-interested claims from 13% with hypothetical stakes to 21% using $10 stakes is not significant. If, however, one includes claims of .9, which are close to pure self-interest, then these claims rise to 37% in treatment 3 whereas the other percentages stay the same. Thus, the percentage of *approximately* self-interested claims with $10 stakes is significantly greater than that with hypothetical stakes but the same as with $5 stakes.

Forsythe, et al. also explored the effects of stakes on a widely tested bargaining game called the *ultimatum game*. In this design, the first mover is given a fixed sum of money and is anonymously matched with a counterpart (the second mover). The first mover proposes a division of the sum to the second, and if the second mover agrees, the money is divided as proposed, whereas if the second rejects the proposal, both receive nothing. The ultimatum game possesses several properties that are very useful for evaluating the behavioral predictions of alternate theories. It is the simplest of bargaining games, since there is only one round of decision-making and only one decision per person, yet it is truly a *game* since the players interact strategically (as opposed to the dictator game, which is a misnomer since only the dictator can influence payoffs). Because of the structure of the ultimatum game, the first mover has more power than the second, but empirical tests suggest this is less than the absolute power in the dictator game. Solving the subgame perfect equilibrium by backward induction, a rational and self-interested second mover would accept the smallest offer, and possibly nothing at all.

Knowing this, the optimal move of the first rational and self-interested mover is to make the minimal offer to the second mover.

The predictions of rational self-interest are extreme, and they are rarely even approximated in the many runs of this experiment. Typically, first movers propose positive amounts, where the mode is equal splits, and second movers sometimes reject small positive offers. Moreover, the rejection rate usually rises, in these studies, as the size of the offer falls. Nevertheless, fairness alone cannot account for these results. A purely fair-minded second mover would reject any proposal other than equal splits, even if the inequality favored him or her, and a fair-minded first mover would only propose equal splits. Thus, this "fairness hypothesis" also produces extreme predictions.

Forsythe, et al. conducted ultimatum games in a between subjects design for the same three levels of stakes as their dictator games. The results are summarized in the second rubric of Table 1. The mean first mover claims do not differ significantly from one another in the ultimatum games (characterized by the absence of superscripts 4, 5 and 6 on those means), indicating, in contrast to the dictator games, that hypothetical stakes yield the same results as actual ones. Comparing claims across the dictator and ultimatum games, however, one does observe differences. Mean claims and the percentage of self-interested claims are significantly less in the paid ultimatum games (5 and 6) than in the paid dictator games (2 and 3), indicating that subjects do respond to the strategic differences in the two games. Subjects do not exhibit such a difference, however, when stakes are merely hypothetical (i.e., 1 versus 4). The closeness of mean claims in the ultimatum game is related to the high percentage of equal claims in that case, in fact, hypothetical first movers appear to underestimate their tendency with real stakes to offer equal splits. Mostly, differences in stakes do not matter for the ultimatum game, although the jump from zero to $10 stakes does significantly raise the percentage of equal split claims and reduce the variance in claims. Mean claims, however, are not significantly affected by changes in already positive stakes.

Pure self-interest and pure fairness produce extreme predictions for the dictator and

ultimatum games. A more reasonable assumption about motivations is that players care about both their own interests as well as fairness (e.g., as Fehr and Schmidt, 1999, propose). This explains the tendency of second movers to reject small offers and of first movers to propose generous splits that, nonetheless, often favor themselves. In the ultimatum game, first mover fairness is reinforced by two considerations: their preference for fairness as well as their self-interested concern that overly selfish claims will be rejected by fair second movers. Thus, it is not surprising that first mover claims are fairer in the ultimatum game than in the dictator game.

The results of the Forsythe, et al. experiments suggest that even small stakes yield results equivalent to those with larger stakes. It might be the case, however, that the stakes typical of most experiments are still insufficient to be indicative of behavior when significant economic rewards are at stake. Hoffman, McCabe and Smith (1996) decided to sink some serious money into examining this question using the ultimatum game with stakes of $10 in one treatment and of $100 in another. The results of their study are reported in the bottom rubric of Table 1. Mean claims by first movers do not differ significantly when stakes are raised from $10 to $100. In fact, these means are not significantly different from those in the paid Forsythe, et al. ultimatum games. The $10 and $100 ultimatum games of Hoffman, et al. are also indistinguishable in terms of the percentage of equal split claims and self-interested claims. The standard deviation is significantly greater with $100 stakes, but even this disappears if one eliminates a single outlier: one observation actually corresponds to pure self-interest with the first mover claiming the whole $100 and the second mover consenting. Collectively, the findings from the studies above imply Results 1 and 2.

Result 1: *For a given incentive structure, even small real stakes generate claims equivalent to those on large real stakes.*

Result 2: *Strategically advantaged agents claim, on average, more than the fair share but less than their "self-interest share," i.e., their optimal claim if all agents were self-interested and rationale*.

Result 1 is reassuring regarding our choice to analyze results based on shares as opposed to

absolute levels. Result 2 suggests a trade-off between fairness and self-interest, which we will explore further in the following section.

3.2 *Effects of Variation in Self-Interest Shares on Claims*

Differences in the claims of first movers in dictator versus ultimatum games in the Forsythe, et al. study suggest that the willingness of people to act on fairness is related to their strategically optimal shares as prescribed by self-interest. The bargaining experiments of Spiegel, Currie, Sonnenschein and Sen (1994) provide more thorough evidence on this relationship and on the trade-off between fairness and self-interest. Among the experiments reported in that paper is one involving two rounds of bargaining over $6 stakes. In the first round, the first mover (denoted F) proposes a division of the $6. If the second mover (denoted S) agrees, the money is divided as proposed. If S rejects the proposal, however, the stakes shrink, and now S can make a proposal to F. If F accepts, the shrunken stakes are divided as proposed, and if not, both players receive nothing.

Table 2 summarizes the stakes and results for this experiment. In Game 1, the first round stakes are $6, but if S rejects the proposal of F, the second round stakes shrink to $0. That is, Game 1 reduces to an ultimatum game, and, as previously discussed, a self-interested F claims the whole amount (or that less 1 cent), which is denoted in the "Self-Interest Share" column of Table 2 as 1.00. In Game 2 the second round share shrinks to $1, and F's optimal claim in the first round is $5, or .833 of the $6 stakes, leaving $1 for S.[1] Games 3 through 6 progressively increase the second round stakes with the optimal first round claims indicated in the Self-Interest Shares column.

Table 2 illustrates that the relationship of actual claims to self-interest shares is (weakly) monotonic, viz., as the first mover's optimal self-interest shares decrease, claims by the first mover fall.[2] The first mover responds, then, to deteriorating bargaining power by reducing claims, consistent with self-interest, leading to this result:

Result 3: *Actual claims are moderated by fairness but also respond to changes in self-interest shares, suggesting a well-behaved trade-off between fairness and self-interest.*

This is also consistent with both the implication of dictator-ultimatum game comparisons above and with the results of Andreoni and Miller (2002), who vary the price of giving in dictator games, and find that the preferences of most agents are well-behaved. Also, as the standard deviations reveal, there is the least dispersion in first mover claims when the self-interest shares of both parties converge to fair shares, summarized as follows:

Result 4: *Variance in claims rises with the degree of conflict between self-interest and fair shares*.

Evidence from numerous dictator games implies people differ with respect to their relative preference for fairness. Given that fact, one can expect that, when the tension between self-interest and fairness is greater, the dispersion of claims will be greater. Thus, the standard deviations of claims are smallest when self-interest and fairness correspond to similar claims (Games 3, 4 and 5) and greatest when they do not (Games 1, 2 and 6). Nevertheless, only in the case of Game 3 does a majority of first movers claim the self-interest share, and in Games 1 and 6, no claims coincide with pure self-interest.

There is a noticeable break in several statistics in comparing the first three games, in which the first mover is strategically advantaged, with the last three, in which he is not. One is the significant difference in mean claims between the first and second three sets of games. Another is the statistically significant difference in the percentage of equal claims between the two groups. A final pattern is that the first three games do not result in disagreement (except for two pairs in Game 1), whereas the last three games are significantly more likely to end in disagreement. The percentages of equal claims and of disagreements for these games are illustrated in Figure 1. Thus, in Games 1 through 3, first movers tend to claim unequal shares that favor them, whereas in Games 4 through 6, they are significantly more likely to claim equal shares, suggesting this result:

Result 5: *When multiple allocation rules are salient, many people employ them in a self-serving fashion, choosing the one that yields a greater share for themselves.*

This replicates with real stakes the finding of Messick and Sentis that advantaged subjects claim

larger shares whereas disadvantaged ones appeal to equality. When the first mover is advantaged, claims of equal splits are unusual, but when self-interest shares call for equal shares or less, around one-half of first movers suddenly claim equal splits.

4. Information and Egocentric Biases

The results of the previous section suggest that information about stakes, self-interest and fairness influence egocentric fairness biases and the willingness to behave fairly. This section examines evidence on fairness biases from studies that, in contrast to the prior studies, specifically vary information about facts relevant to self-interest and fairness. We first review and analyze the studies of Roth and his collaborators, which explore how varying information relevant to fairness affects fair behavior and outcomes. Then we examine studies that manipulate stakes or information about stakes, which confirm egocentric fairness biases and suggest that they can be traced to self-deception, i.e., to the tendency of people to process information in a self-serving manner.

4.1 *Variation in Information Relevant to Self-Interest and Fairness*

Roth and his colleagues Malouf and Murnighan conducted a series of bargaining experiments that are especially helpful for illuminating the relationship between information and fairness. The design of these studies shares a number of common features: Subjects are paired anonymously, but bargaining is much less structured than in the previously discussed experiments.[3] A distinctive feature of this design is that subjects do not bargain directly over stakes but rather over lottery tickets that determine their probability of winning a prize. That is, a subject with, say, 60% of the tickets has a 60% probability of winning his prize. In almost all treatments, these prizes differed across pairs, e.g., one player had a $3 prize whereas the other player had a $9 prize. These studies, then, precluded the equal outcomes generally considered fair in such experiments. Players could, at best, achieve equal expected values by allocating fewer tickets to the player with the high prize. This would occur, for example, if a $3 player received 75% of the tickets and a $9 player 25%, creating equal expected values of $2.25. Thus, equality in the variable being allocated (tickets) corresponds here to a different allocation from

equality in the variable players presumably value (expected earnings). Subjects always knew their own prizes, and information was varied across treatments in these experiments in terms of what subjects knew about the value of their counterparts' prizes.

In Roth and Malouf (1979), players bargained over lottery tickets under one of two conditions. In the *full information* condition, both players knew the values of both prizes. In the *partial information* condition, each player knew only his or her own prize, but not that of the counterpart. Subjects played four games in random order under just one of the two information conditions. This is summarized in the top two rubrics of Table 3. In Games 1 and 2, both players had equal prizes of $1.00, whereas in Games 3 and 4, one player had a prize of $1.25 (the Low Prize Player) and the other of $3.75 (the High Prize Player). Games 1 and 2 (3 and 4, respectively) differed in inconsequential ways for our purposes, so their results are collapsed in Table 3.[4]

Treatments 1 and 2 in Table 3 deal with Games 1 and 2, where both players had equal $1.00 prizes. Taking the fair division of tickets to be the one that produces equal expected payoffs, this gives each player, of course, one-half of the tickets, indicated in the Fair column (for consistency, all results in Table 3 are expressed for the High Prize Player, although, in this case, there is no difference). The mean share of tickets actually agreed upon does not differ significantly from .5 in either information condition. In fact, the two information conditions also do not differ significantly in terms of mean, variance, percentage of equal and fair shares (which are the same, in this case), or percentage of bargaining encounters that led to disagreement.[5] These experiments lack the explicit asymmetry in bargaining power present in the previously discussed games, and one finds that a high percentage of outcomes are fair. The increased information in treatment 2 versus treatment 1 appears to have mostly beneficial effects by increasing the fraction of fair agreements and reducing disagreements (although variance also rises slightly), but these values are already so favorable even in the partial information condition that these improvements are not significant. Since there is no tension between allocating tickets equally and equalizing expected values, however, these treatments are not useful for examining

information effects when there are conflicting fairness rules. Therefore, in all of the discussion that follows, I will refer only to the results for the other treatments by Roth and his colleagues in which subjects may appeal to differing rules.

The unequal prizes in treatments 3 and 4 of Roth and Malouf imply that expected payoffs are equalized when the High Prize Player receives 25% of the tickets. When each player only knows his own prize in the partial information condition, however, mean shares do not differ significantly from .5, and 69% of agreements result in equal splits of tickets. If both players are fully informed about prize values, however, there is a significant shift in mean ticket shares toward fair shares, as evidenced in Figure 2. The percentage of agreements with equal shares drops from 69% to only 17% while the percentage with fair shares rises from 0% to 22%, both significantly. On the other hand, increasing information does not appear to aid convergence, as the standard deviation rises in treatment 4 versus 3 (and 1 and 2), and the percentage disagreements rises, although not significantly. We will see that precisely these patterns of effects from increasing information repeat themselves in other studies.

Roth, Malouf and Murnighan (1981) built on this work by increasing the average stakes, the number of different prize values, and the number of levels of information. In the *high information* condition, subjects were able to calculate fair shares, whereas the information in the *intermediate* and *low information* conditions did not permit them to do so.[6] Since mean shares in the latter two conditions did not differ significantly, we combine them. The results of this experiment are summarized in the middle rubrics of Table 3. Each subject played the high and the low prize position in Games 1 and 3, but since the prizes were the same for these games, these results have been consolidated. Note that Games 2 and 4 of this study require more extreme differences in ticket shares to equalize expected values, as can be seen from the Fair column. Comparing the Low/Intermediate and High information conditions for each treatment, one sees that increasing information causes mean shares to move toward equalizing expected values, variance to increase, percentage equal shares to fall, percentage fair shares to rise (except treatments 3 and 4, which are the same) and percentage disagreements to rise. These are the same

patterns observed in the Roth and Malouf study and are generally significant at the 5% level for means and standard deviations.

The procedures of Roth and Murnighan (1982) were similar to those of Roth and Malouf (1979), with a few crucial differences. The prizes were always equal to $5 and $20, and there were four information conditions with respect to what each subject knew about the other's prize: 1. Neither player knows his opponent's prize, 2. Only the player with the $20 prize knows the other's ($5) prize, 3. Only the $5 player knows the other's ($20) prize, and 4. Both players know each other's prize.[7] Considering the results in the bottom rubric of Table 3 for each of the information treatments progressively from 1 to 4, the mean share moves away from equality and toward fair shares, as the percentage of equal shares falls. Nevertheless, the tests of significance point toward a break between the first and second two sets of treatments. The mean share when only the $20 player knows the other's prize does not differ significantly from the case in which neither knows. But when the $5 player knows the other's prize (i.e., that it is $20), there is a significant shift toward equal expected payoffs. Moreover, this is not significantly different from the case in which both know both prizes. This suggests that providing information relevant to fairness does not strongly motivate the advantaged player to act on it, but it does cause the disadvantaged player to work toward fairer outcomes. In fact, the lack of significant differences between treatments 3 and 4 by any measure suggests that fair outcomes in this study are due solely to the efforts of the disadvantaged party to defend its fair share. This is consistent with evidence Straub and Murnighan (1995) provide from ultimatum games that first movers make significantly lower offers when the second movers do not know the size of the stakes and are less likely, therefore, to reject small and unfair shares. In terms of convergence, the only significant difference in variances is between treatment 1 and all others. Also, the only significant difference in the rate of disagreements is between treatment 1 and one other (treatment 3). The implication of these results is that information that introduces alternate fairness rules (equal expected values) increases dispersion, but additional information has no significant impact on convergence.

The general conclusions one can draw from the studies of Roth and his colleagues are the

following. Increased information tends to reduce bias and sometimes to increase the percentage of fair shares, at least, if the information is sufficient to calculate fair outcomes. On the other hand, additional information seems to have an unfavorable effect on convergence of shares: standard deviations of shares rise as do rates of disagreement, although the latter is usually not significant. The increased variance is associated with information that introduces conflicting allocation rules.

The foregoing result on convergence, however, is misleading: if parties genuinely value fair outcomes, here fair expected shares of earnings, then shares of tickets are merely an intermediate means of achieving the more important goal. In that case, convergence should be evaluated in terms of the earnings, or expected earnings, that the bargaining process produces. These studies do not report the actual payoffs resulting for the randomly chosen winners, but one can easily calculate the distribution of expected payoffs from the product for each player of each his share of tickets and his prize. The results of this exercise for each of the treatments are reported in Table 4. One sees that the mean expected payoff to the High Prize Player always moves toward equality as information rises, consistent with the pattern of change in the shares of tickets. Calculating the expected payoffs to both players, the standard deviation always falls, sometimes significantly, with increased information. Figure 3 illustrates these results for treatments 3 and 4 of Roth and Malouf.

Although greater information is typically associated with an (insignificant) increase in the rate of disagreement in the studies above, it can also reduce disagreements, as demonstrated by the study of Camerer and Loewenstein (1993). They conducted ultimatum games with five levels of stakes: $1, $3, $5, $7 and $9. In the *certain amount* treatment, first movers indicated how much they would offer, and second movers indicated the minimum offers they were willing to accept, in both cases for each of the five stakes. Subjects were matched, one decision was randomly chosen for payment, and if the first mover's offer was no less than the second mover's minimum acceptable offer, the offer was paid out; otherwise both received nothing. The *uncertain amount* treatment, conducted with a separate group of subjects, was the same, except

that only first movers knew the stakes and made five separate decisions for each case; the second movers stated a single minimum acceptable offer without knowing the stakes. In the certain amount condition, the means and standard deviations of first mover offers (calculated as shares of the total stakes) were very similar across the five stakes as were means and standard deviations of the second mover minimum acceptable offers. Moreover, the disagreement rates for all levels of stakes were very close to the overall average of 15%. In the uncertain amount condition, the means and standard deviations of first mover offers were again quite similar to one another as well as to those in the certain amount condition. But, since second movers did not know what level of stakes their minimum accept offer applied to, higher rates of disagreement resulted that, across all stakes, averaged 39%.

Thus, we conclude the following. In the face of countervailing bargaining power, the overall effect of increased information is to reduce bias in the claims settled (Table 3, Figure 2) as well as to reduce bias and dispersion in the outcomes, or expected outcomes, that people value (Table 4, Figure 3). Although information can supply fodder for conflicting claims, it is beneficial (especially in the hands of the disadvantaged person) if it is relevant to fairness, i.e., information that is necessary to calculate fair allocations according to accepted norms. Result 6 summarizes this.

Result 6: *Under symmetric bargaining, providing information relevant to fairness results in fairer mean expected outcomes, a higher percentage of fair expected outcomes and greater convergence (i.e., lower variance in expected outcomes). Disagreement rates can be favorably or unfavorably affected, depending on the structure of bargaining.*

4.2 *Egocentric Processing of Information*

We have seen evidence from numerous studies that information affects the egocentric fairness bias, which in turn can affect claims, allocations and disagreement rates. We turn now to studies that explore the link between information and bias and that indicate this bias is due to self-serving processing of information. These studies also provide a more realistic set of conditions than most laboratory experiments, including contextual richness and fair allocations

that, as in most real world contexts, do not necessarily correspond to equal splits.

Babcock, Loewenstein, Issacharoff and Camerer (1995) created a contextually rich laboratory experiment, in which each subject was first assigned to the role of plaintiff or defendant in a legal dispute.[8] Subjects read 27 pages of testimony abstracted from an actual case and were informed that a judge had read the same materials and had decided how much, if anything, to award to the plaintiff. Then, they provided their judgments of 1) what they considered a fair settlement, and 2) their best estimate of what the judge's award would be. Finally, the parties had 30 minutes to reach an agreement (the division of the experimental stakes of $10 were in proportion to the agreement), or the judge's decision was imposed, less legal costs.[9] This describes the condition with *ex ante stakes*, where each subject knew his role in the dispute at the start. There was a second condition with *ex post stakes*, which was identical, except that subjects were not informed whether they were the plaintiff or the defendant until just before beginning negotiations, but after they had read the court materials and given their estimates of the fair settlement and the judge's award. The idea is that egocentric processing of information begins when roles are known, and that manipulation of the timing of that information produces stable biased beliefs, when roles are known ex ante, and stable unbiased beliefs, when roles are revealed ex post. Thus, Babcock, et al. predicted that the ex ante treatment, in comparison to the ex post treatment, would result in more biased judgments of the fair settlement and the judge's award and lead to a higher rate of disagreement.

The results of this experiment are presented at the top of Table 5. The high stakeholder refers to the advantaged player, which is the defendant in this experiment. This is consistent with the judge's award of .306 to the plaintiff, which leaves .694 for the defendant. One sees dramatic evidence of fairness bias in the results for the condition with ex ante stakes. The high stakeholder has significantly higher estimates of what he should receive according to the fair judgment (.791) and judge's award (.761) than the low stakeholder (.593 and .575, respectively). In the ex post stakes condition, these two estimates fall significantly for the high stakeholder (to .690 and .638) and rise significantly for the low stakeholder (to .753 and .707). The differences between high

and low stakeholder estimates are significantly greater in the ex ante condition than in the ex post condition. Moreover, the ex post condition eliminates any bias in fairness judgments at conventional levels of significance. Finally, in comparing the ex ante and ex post agreements, the difference in agreed shares is borderline significant and the difference in disagreement rates is highly significant.

All of the Babcock, et al. predictions are borne out: parties in the ex ante condition have more biased fairness judgments and are much less likely to reach a settlement. In addition, differences in fairness judgments were a significant (directly related) predictor of the time parties took to settle and the probability of impasse. The results of this study are summarized below:

Result 7: *The self-serving fairness bias of stakeholders is due to egocentric processing of information. This bias is an important and significant predictor of bargaining impasse*.

It is reasonable to ask whether the powerful findings of this seminal study are robust with respect to changes in experimental procedures.

Gächter and Riedl (2001) constructed a bargaining experiment, similar in several ways to Babcock, et al., in which subjects were randomly and anonymously paired and asked to place themselves in the role of one of two department heads in a firm. The subjects were then told that, historically, the department head who scored better on a general knowledge quiz also performed better as head and, therefore, received twice the salary of the lower scoring head. Next, subjects took the quiz, were informed which of the two bargaining partners scored higher, and responded to a question about what they considered to be the fair distribution according to a neutral arbitrator. They then had 15 minutes to reach an agreement over 205 Austrian shillings (about US $18) through free-form bargaining or receive zero payoffs, if no agreement was reached. This was their condition with *ex ante stakes*. The *ex post stakes* treatment was identical, except that the fairness question was posed right after reading the instructions but before subjects knew whether they had scored higher on the test, indeed, before they even took the test.

Gächter and Riedl find that fairness judgments correlate significantly with a number of behaviors, including opening offers, bargaining duration, concessionary behavior and agreed

shares. They also report an important *entitlement effect*, i.e., both parties respond strongly to the two-to-one salary frame of the experiment, which suggests the winner of the quiz is entitled to two-thirds of the earnings. The bottom of Table 5 summarizes this result, where the high stakeholder in this study is the winner of the quiz. In the ex ante stakes treatment, the fair share to the winner according to both high and low stakeholders is close to two-thirds. Nevertheless, as this table shows, the results on fairness bias and disagreement rates from this experiment are not as large or significant as in the Babcock, et al. study. In the ex ante treatment, high stakeholders consider their fair share to be greater than do low stakeholders, but this difference (.024) is not significant at conventional levels (p=.08). The difference in fair judgments between high and low stakeholders in the ex ante treatment is significantly different, however, from that in the ex post treatment. Gächter and Riedl also report opening offers in the negotiations and find that these differ significantly for high and low stakeholders in both treatments. High stakeholders lower and low stakeholders raise their opening offers slightly in the ex post treatment in comparison to the ex ante treatment, although these differences are not significant. Thus, self-interest is still at play in this study but not so prominently because of an egocentric bias. Finally, the disagreement rate is no higher (indeed, is lower) in the ex ante condition in comparison to the ex post condition, and the authors find no evidence of a direct effect of fairness judgments on disagreement rates.

The Gächter and Riedl experiment finds more modest evidence of an egocentric fairness bias and its impact on bargaining than the Babcock, et al. study. In comparing the two studies, two important procedural issues stand out that might explain these differences. First, Babcock, et al. present their subjects with a much more lengthy list of facts than Gächter and Riedl. This provides a rich basis for choosing, interpreting, weighting and recalling facts in a self-serving fashion. As Dunning, Meyerowitz and Holzberg (1989) report, self-serving evaluations rise with an increase in the number of criteria at one's disposal. In an earlier study employing mostly the same procedures as the Babcock, et al. ex ante treatment, Loewenstein, Issacharoff, Camerer and Babcock (1993) asked parties after negotiation to list and rate the importance of arguments

favoring themselves and their partners. Parties were significantly more likely to recall more arguments favoring themselves and to rate self-serving arguments as more important. Second, Gächter and Riedl explicitly present subjects with a statement of fair shares, whereas Babcock, et al. do not. As Gächter and Riedl themselves conjecture (pg. 19), this precedent apparently conveyed a strong sense of entitlement, even to the disadvantaged low stakeholders, who usually judged their own fair share as less than one-half. Comparing the various studies, we come to the following conclusion:

Result 8: *The egocentric fairness bias of stakeholders is enhanced by more ambiguous facts and is diminished by an independently determined estimate of fair shares*.

These studies represent significant contributions to the literature on egocentric fairness biases and stimulate a number of unresolved questions. They rely on self-reports of fairness beliefs to measure biases, but these do not directly determine any actual allocations. In addition, their context is bargaining, which is appropriate for the types of situations their authors wish to examine, but the question remains whether fairness biases arise even when strategic interaction is not involved. For the purpose of reducing biases, withholding information about roles, although a useful mechanism for identifying biases experimentally, does not represent an effective means of alleviating biases in most real world situations where parties know their stakes, as these authors acknowledge. Finally, all of the foregoing experimental studies analyzed the behavior of individuals who were in the position of stakeholder. The last two studies addressed bias by manipulating the timing of information about roles. Another approach to identifying biased and unbiased beliefs about fairness is to manipulate the roles themselves.

In a set of dictator games (Konow, 2000), I sought to address these and certain other questions related to justice. In this experiment, subjects first generated earnings jointly by preparing letters for a mass mailing, after which one randomly assigned subject (the dictator) allocated the joint earnings from the first phase. In one set of treatments, the "discretionary differences" condition, 50 cents is credited to the pair for each letter prepared, and the percentage of joint earnings attributable to one of the paired subjects varied because of differences across

subjects in the number of letters prepared in the allotted time. There were two treatments in the discretionary differences condition. In the *Stakeholder* treatment, the dictator was one of the subjects in the pair, whereas in the *Spectator* treatment,[10] the dictator was a subject from a third group, who did not prepare letters but was paid a fixed fee to allocate the joint earnings to each of two other subjects who did prepare letters. The prediction of the theory outlined in that paper was that spectators would allocate in proportion to the relative contributions of the subjects to joint earnings (in accordance with equity theory, e.g., Homans, 1958, Adams, 1965, and Walster, Walster and Berscheid, 1973), whereas stakeholders were predicted to take an amount equal to or greater than the proportional amount.

The results for these treatments are summarized in the top rubric of Table 6. In the first treatment, contributions by one group averaged .491, which is the average fair allocation to that group. The actual allocations of third party dictators to that group averaged .494, an insignificant difference. Simply analyzing differences in means, however, does not evaluate the prediction of proportionality, only that allocations are fair, on average. So a paired difference test of contributions to allocations was conducted, and it establishes that the allocations of spectators do not differ significantly from fair (proportional) allocations ($p=.88$).[11] Turning to the second treatment under the discretionary differences condition, stakeholders allocated to themselves .644 of the earnings, an amount that differs significantly from the self-interest share of 1.0 and from their fair share according to average contributions of .516, both by a test of differences in means as well as by the paired difference test. Finally, an examination of the standard deviation of actual allocations reveals that variance is significantly lower among spectators, indicating that this kind of impartiality not only eliminates bias but is also associated with convergence.

A different set of treatments generated similar results with respect to stakeholders versus spectators while providing additional information about motivation. In the “exogenous differences” treatments, all subjects had enough time to prepare an equal number of envelopes, but the earnings due to each subject differed because of arbitrary differences in the per letter credits across the two subjects in each pair, always to the advantage of the dictator. According to

the theory in that study, fair allocations in this condition were always equal splits.[12] As in the discretionary differences condition, there were Stakeholder and Spectator treatments for the exogenous differences condition, as shown in the bottom rubric of Table 6. The results are also illustrated in Figure 4. In treatment 1, spectators allocated .508 to one group, which differs insignificantly from the fair share of .5.[13] The stakeholders in treatment 2, on the other hand, allocated .592 of the earnings to themselves, significantly more than .5. The standard deviation of spectators is also significantly less than that of stakeholders, as in the discretionary differences condition. The results from both conditions, then, confirm the bias of stakeholders, and the impartiality and convergence of spectators.

An additional goal of this study was to identify and quantify as cleanly as possible the sources of unfairness when real stakes are involved. Unfair behavior was decomposed into two parts: the kind of *unadulaterated self-interest*, or deliberately selfish behavior, that appears to dominate in the Gäechter and Riedl study, and the Babcock, et al. finding of apparent *self-deception*, or unfairness owing to the fact that people sometimes deceive themselves into believing it is fair to be unfair.[14] The stakeholder treatment in the exogenous differences condition contained a second decision designed for this purpose. After choosing their own divisions, dictators were then informed for the first time that they would be making an additional decision. They were also to allocate earnings between two other subjects, whose unequal earnings from the first phase were the same as the unequal ones of the dictator and his/her counterpart. Although fairness calls for equal shares, to the extent the dictators had convinced themselves that it was fair to take more than one-half in their own case, they should now allocate more to their counterpart with higher credits in the new pair. In fact, standard dictators gave .558 of the earnings to their counterpart in the new pair, in comparison to the .592 they took for themselves in the earlier round. That is, around 60% of their unfairness (allocations to themselves in excess of .5) appears to be due to self-deception, whereas the remaining 40% is unadulterated self-interest.[15] The foregoing findings are summarized below:

Result 9: *The unfairness of stakeholders is due mostly to self-deception, i.e., beliefs about*

fairness that are egocentrically biased. The views of spectators, on the other hand, are unbiased and converge significantly more than those of stakeholders.

5. Conclusions

The analysis of this paper points toward the following conclusions. Even small real stakes produce measures of bias and convergence in claims that are similar to those with large stakes. There is a well-behaved trade-off between fairness and self-interest. Claims become more biased and disperse and disagreements more frequent as the fair claim increasingly conflicts with self-interest and with the introduction of information that provides a basis for competing allocation rules. Biased claims can be traced to two sources: pure self-interest and self-deception arising from egocentric processing of information. If bargaining power is balanced, however, the overall effect of increased information is favorable: settled claims are less biased, and outcomes are less biased and less disperse, although disagreement rates can increase or decrease, depending on the structure of bargaining. Bias and variance can be reduced by addressing the impact of personal stakes, for example, by the timing of information about roles, by an independent statement of fair shares, or by removing stakes altogether and placing individuals in the role of spectator.

These conclusions differ considerably from the Rawlsian "conventional wisdom" by indicating that justice is better understood and better realized by adding rather than removing information. This is promising for the feasibility of fair constitutional contracts, laws and institutions, and suggests ways of resolving conflicts. Operating under realistic information conditions, these findings raise the prospect of employing empirical methods to identify impartial justice, which could help address important questions of interest in philosophy, law and the social sciences. Such methods could inform political discourse by helping to rule out biased claims that are erroneously defended based on self-serving manipulation of justice principles. As the evidence shows, however, stakeholders develop blind spots about justice that are particularly insidious because of the tendency to nurture self-serving beliefs about it. It can be difficult to dislodge such biases, because people believe they acting on principle. In light of this, I conclude with some observations about the prospects for impartial justice.

A direct application of one empirical lesson is to use the fairness judgments of impartial spectators, i.e., groups of informed individuals who have no stake in the matter under consideration. In fact, that seems to be the concept around which an independent judiciary is designed. Judges and juries are supposed to be unbiased adjudicators to whom the relevant facts are presented so that they may decide the just distribution of benefits or burdens. Of course, the costs of legal proceedings prohibit this as a means to resolve all conflicts, and legal recourse is, at best, an imperfect approach. More importantly, although third party judgments can prove a very useful tool for social science investigation, one would be seriously mistaken to conclude that they are necessarily the best apparatus for social and political decision-making. Instead, it is better to think of impartial spectatorship as did Smith: individuals should appeal to and be guided by their own impartial spectator. We have seen how fairness can emerge when conflicting parties engage in symmetric bargaining and when they are provided with an independent estimate of fair shares. Other studies provide additional evidence on how fairness and impartiality might be achieved in individual and collective decision-making.

Babcock and Loewenstein (1997) report the results of various interventions aimed at reducing or eliminating the egocentric fairness bias in their experiments. They found informing subjects of the bias had no effect, either on the discrepancy in fairness assessments and predictions of the judge's decision or on the likelihood of settlement. Having subjects write an essay arguing the opponent's case was similarly ineffectual. They did find, however, that informing subjects of the egocentric bias *and* having them list weaknesses in their own case significantly reduced the differences in predictions of the judge's decision and lowered the rate of impasse from 36% to only 4%. Oberholzer-Gee, Bohnet and Frey (1997) describe the siting of nuclear waste facilities, and their findings support the importance of impartiality, fairness and information. They report a willingness, uncharacteristic in such cases, on the part of a Swiss community to accept the siting, which the authors trace to fair, efficient and informed decision-making procedures. The promise of informed and impartial justice will lead, I hope, to further investigations of mechanisms to resolve conflicts and will shed light on collective action.

Notes

[1] Solving the subgame perfect equilibrium by backward induction, note that the second round is an ultimatum game with $1 stakes: S can offer F 1 cent, or nothing, and F will accept. In the first round, then, F knows this and, therefore, offers $1 to S in the first round so that F can secure $5 in the first round rather than nothing in the second.
[2] Claims in Games 1 and 2 are the same: perhaps the failure of Game 1 claims to be greater is a reflection of F's caution given the finality of this game in comparison to the other ones in which something can still be salvaged if S initially rejects.
[3] Specifically, players were given 10 or 12 minutes, depending on the study, to communicate and come to an agreement via computers, or to earn nothing. Any messages were allowed, except those that identified the subject, proposals were binding on the sender, and bargaining ended when a subject accepted a sender's proposal or when time ran out.
[4] Games 2 and 4 contained the additional restriction that the High Prize Player was not permitted to receive more than 60% of the tickets. Although this constraint would have been binding on subjects pursuing equal expected values if it had been placed on Low Prize Players, it was not so when placed on High Prize Players. Moreover, the only statistically significant difference Roth and Malouf report here is between Games 3 and 4 in the partial information condition.
[5] Note that the Mean, Standard Deviation, % Equal and % Fair are for those pairs that reached an agreement, i.e., the first number in the N column, whereas the % Disagree is out of the total sample, i.e., the second number in the N column.
[6] Specifically, each subject played four games with differing prizes under just one of the three information conditions. As in Roth and Malouf, players bargained for lottery tickets, but here the tickets were expressed in terms of an intermediate commodity, viz., the prize was a certain number of chips that had a certain monetary value. The number of chips, the value per chip and the total prize (the product of the two previous values) differed across the four games. In every treatment, each player knew his own number of chips, values per chip and total prize, but the information each player was given about his opponent's prize was varied across conditions. In the *high information* condition, each player knew the number and value of chips in his opponent's prize. In the *intermediate information* condition, each player knew the number of chips in his opponent's prize, but not the value per chip. In the *low information* condition, neither player knew the number or value of chips in his opponent's prize.
[7] There was one other treatment variable consisting of two conditions pertaining to common knowledge, and the design was between subjects: each of the eight (4 information × 2 common knowledge) treatments was conducted with a different set of subjects. Under *common knowledge*, both players are told what information is available to whom. For example, under common knowledge, both the $20 and $5 players in treatment 2 above know that the $20 player knows both prizes but that the $5 player only knows his own prize. Under *non-common knowledge*, it is still the case that the $20 player knows both prizes and the $5 player knows only his own prize, but both are told that the other may or may not know their prize. In any case, mean shares do not differ significantly across common knowledge conditions for any of the four information treatments, so these results are combined.
[8] Babcock, et al. introduced an innovative approach to analyzing bias in bargaining experiments, making a break with two literatures. Previous theoretical research in law and economics suggested that cases fail to settle out of court because of random errors in litigants' estimates of the value of going to trial, which implies that increasing information to the parties should reduce disputes. Previous experimental research on bargaining usually strived for contextual simplicity, often reducing or eliminating information about specifics of the persons and conditions involved.
[9] Subjects were paid a bonus at the end of the experiment for accurately estimating the latter value within a range. The plaintiff was suing for $100,000, and the settlement could lie anywhere between that and $0, so the defendant's initial experimental endowment of $10 was allocated proportionately at a ratio of 1:10,000.
[10] In the paper, these were called the "Standard" and "Benevolent" Dictator treatments, respectively.
[11] In addition, a regression of allocations on earnings indicates a slope of one, intercept of zero and R^2 of 0.63, and the fraction of benevolent dictators allocating exactly in proportion to earnings was 46%. In fact, correcting for an experimental artifact, the R^2 rises to 0.98 and the percentage of exactly proportional allocations to 79%. This is powerful corroboration of a strong prediction of equity theory for the case of spectators. Indeed, this was the first such general confirmation of the exact proportionality rule, to this author's knowledge. Proportionality has surfaced, though, in subsequent work: Ellingsen and Johannesson (2002) combine produced earnings, as in this study, with manipulation of information about roles, as in Babcock, et al., and also find evidence that perceived contributions are proportional to first mover claims in their ex post treatment.

[12] Specifically, the average credit per pair was still 50 cents, but if the dictator's credit was, say, 75 cents per letter, then the other subject's credit was 25 cents per letter. Thus, although monetary contributions differed, each subject produced one-half of the letters in his or her pair. This feature was designed to test a prediction of the theory presented in that paper that fair allocations are proportional only to the "discretionary" variables agents control (e.g., letters produced in the experiment) but are unrelated to the "exogenous" variables agents do not control (e.g., per letter credits), even if the uncontrolled variables affect contributions (e.g., earnings). This provides a criterion for distinguishing among the variables that are relevant for applying proportionality, a distinction most of equity theory did not make.

[13] In fact, 87% of spectators split the earnings exactly in halves. Note also that the paired difference test in this case where fair allocations are always .5 is the same as a standard difference of means test.

[14] Although most researchers seem to mean the latter when speaking of an egocentric bias, unfair biases actually have these related but separate origins. This decomposition is an application of Festinger's theory of cognitive dissonance (1957): when agents experience displeasure from conflicting facts, they can respond by trying to alter the facts or by changing their beliefs about them. In the context of justice, Kay and Jost (2003) point out that people confronted with such situations (e.g., fairness is good, the world is not fair) can respond one of two ways: they can try to rectify or rationalize the injustice. In the context of the dictator experiment, the dictator can allocate some amount to the counterpart and/or nurture self-serving beliefs about what is fair. As Hafer (2000) notes, self-reported beliefs in these cases are suspect for a number of reasons. The dictator experiment, on the other hand, by employing real monetary allocations in a non-strategic environment, provides incentives that are compatible with truthful revelations of one's beliefs.

[15] In fact, there are several reasons to believe that this is a conservative estimate of the effect of self-deception. One is that this experiment is contextually simple and provides little basis for egocentric information processing. In addition, given the anonymity of this design, some dictators might believe, or choose to believe, that there are not really any counterparts. I found anecdotal evidence of this in responses to post-experimental questionnaires, and Frohlich, Oppenheimer and Moore (2001) found systematic evidence of such doubts in dictator games. In the current experiment, these doubts lower estimates of self-deception by making it cheap to be selfish in the first round and fair in the second round. The fact that such a large share of unfairness can be traced to self-deception despite these concerns underscores both the importance and the difficulty of achieving impartiality as a stakeholder. On the other hand, the spectator treatments of this experiment indicate how, when personal stakes are removed, bias vanishes and a very high level of consensus is achieved.

References

Adams, J. S. (1965). "Inequality in Social Exchange" Leonard Berkowitz, ed., Advances in Experimental Social Psychology 2, New York, London: Academic Press Inc., 267-299.

Andreoni, J. and Miller, J. (2002). "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism," *Econometrica*, *70*(2), 737-753.

Babcock, L. and Loewenstein, G. (1997). "Explaining Bargaining Impasse: The Role of Self-Serving Biases," *Journal of Economic Perspectives*, *11*(1), 109-126.

Babcock, L., Loewenstein, G., Issacharoff, S. and Camerer, C. (1995). "Biased Judgments of Fairness in Bargaining," *American Economic Review*, *85*(5), 1337-1343.

Babcock, L., Wang, X. and Loewenstein, G. (1996). "Choosing the Wrong Pond: Social Comparisons in Negotiations That Reflect a Self-Serving Bias," *The Quarterly Journal of Economics*, *111*(1), 1-19.

Buchanan, J. M. (1985). *Liberty, Market and State*, New York: New York University Press.

Camerer, C. and Loewenstein, G. (1993). "Information, Fairness, and Efficiency in Bargaining," in Barbara A. Mellers and Jonathan Baron, eds., *Psychological Perspectives on Justice*, Cambridge, MA: Cambridge University Press, 155-179.

Cross, K. P. (1977). "Not *Can*, but *will* College Teaching Be Improved?" *New Directions for Higher Education*, Spring, *17*, 1977, 1-15.

Dunning, D., Meyerowitz, J. A. and Holzberg, A. D. (1989). "Ambiguity and Self-Evaluation: The Role of Idiosyncratic Trait Definitions in Self-Serving Assessments of Ability," *Journal of Personality and Social Psychology*, *57*(6), 1082-1090.

Ellingsen, T. and Johannesson, M. (2002). "Does Impartial Deliberation Breed Fair Behavior? An Experimental Test," University of Stockholm: Economics, December.

Fehr, E. and Schmidt, K.M. (1999). "A Theory of Fairness, Competition and Cooperation," *Quarterly Journal of Economics*, *114*(3), 817-868.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*, Stanford: Stanford University Press.
Forsythe, R., Horowitz, J.L., Savin, N.E. and Sefton, M. (1994). "Fairness in Simple Bargaining Experiments," *Games and Economic Behavior*, May, *6*, 1994, 347-369.
Frohlich, N., Oppenheimer, J. and Moore, B. (2001). "Some Doubts About Measuring Self-Interest Using Dictator Experiments: The Costs of Anonymity," *Journal of Economic Behavior and Organization*, *46*(3), 271-290.
Gaertner, W., Jungeilges, J. and Neck, R. (2001). "Cross-Cultural Equity Evaluations: A Questionnaire-Experimental Approach," *European Economic Review*, *45*, 953-963.
Gächter, S. and Riedl, A. (2001). "Moral Property Rights in Bargaining," University of Amsterdam, University of St. Gallen.
Gorman, R.F. and Kehr, J.B. (1992). "Fairness as a Constraint on Profit Seeking: Comment," *American Economic Review*, *82*(1), 355-358.
Hafer, C.L. (2000). "Do Innocent Victims Threaten the Belief in a Just World? Evidence from a Modified Stroop Task," *Attitudes and Social Cognition*, *79*(2), 165-173.
Homans, G.C. (1958). "Social Behavior as Exchange." *American Journal of Sociology* 63, pp. 597-606.
Kahneman, D., Knetsch, J.L. and Thaler, R. (1986). "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *American Economic Review*, September *76*, 1986, 728-741.
Kay, A.C. and Jost, J.T. (2003). "Complementary Justice: Effects of "Poor but Happy" and "Poor but Honest" Stereotype Exemplars on System Justification and Implicit Activation of the Justice Motive," *Journal of Personality and Social Psychology*, *85*(5), 823-837.
Kirchgässner, G. (2004). "On Minimal Morals," manuscript, University of St. Gallen.
Kolm, S.-C. (1996). "Moral Public Choice." *Public Choice* 87:1-2, pp. 117-141.
Konow, J. (2000). "Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions," *American Economic Review*, *90*(4), September 2000, 1072-1092.
Konow, J. (2003). "Which is the Fairest One of All?: A Positive Analysis of Justice Theories," *Journal of Economic Literature*, *41*(4), December, 1186-1237.
Loewenstein, G., Issacharoff, S., Camerer, C. and Babcock, L. (1993). "Self-Serving Assessments of Fairness and Pretrial Bargaining," *Journal of Legal Studies*, *22*(1), 135-159.
Messick, D.M. and Sentis, K.P. (1983). "Fairness, Preference, and Fairness Biases," in D.M. Messick and K. Cook, eds., *Equity Theory: Psychological and Sociological Perspectives*, New York: Praeger Publishers, 61-94.
Oberholzer-Gee, F., Bohnet, I., and Frey, B. (1997). "Fairness and Competence in Democratic Decisions." *Public Choice* 91:1, pp. 89-105.
Rawls, J. (1971). *A Theory of Justice*, Cambridge: Belknap Press of Harvard University Press.
Roth, A.E. and Malouf, M.W.K. (1979). "Game-Theoretic Models and the Role of Information in Bargaining," *Psychological Review*, *86*(6), 574-594.
Roth, A.E., Malouf, M.W.K. and Murnighan, J.K. (1981). "Sociological Versus Strategic Factors in Bargaining," *Journal of Economic Behavior and Organization*, June, *2*, 1981, 153-177.
Roth, A.E. and Murnighan, J.K. (1982). "The Role of Information in Bargaining: An Experimental Study," *Econometrica*, *50*(5), 1123-1142.
Rowley, C.K. (1993). "Justice." In C.K. Rowley, ed., *Social Choice Theory, Volume III*, Hants: Edward Elgar Publishing, Limited, pp. 185-210.
Schokkaert, E. and Capeau, B. (1991). "Interindividual Differences in Opinions About Distributive Justice," *Kyklos*, *44*(3), 325-345.
Smith, A. (1759) [1809]. *The Theory of Moral Sentiments*, Glasgow: R. Chapman.
Smith, A. (1776) [1977]. *The Wealth of Nations*, London: J. M. Dent & Sons.
Spiegel, M., Currie, J., Sonnenschein, H. and Sen, A. (1994). "Understanding When Agents Are Fairmen or Gamesmen," *Games and Economic Behavior*, *7*(1), 104-115.
Straub, P.G. and Murnighan, J.K. (1995). "An Experimental Investigation of Ultimatum Games: Information, Fairness, Expectations, and Lowest Acceptable Offers," *Journal of Economic Behavior and Organization*, *27*, 345-364.
Svenson, O. (1981). "Are We All Less Risky and More Skillful Than Our Fellow Drivers?" *Acta Psychologica*, February, *47*, 1981, 143-148.
Van Dijk, F., Sonnemans, J., and van Winden, F. (2002). "Social Ties in a Public Good Experiment," *Journal of Public Economics*, *85*(2), 275-299.
Walster, E., Walster, G.W. and Berscheid, E. (1973). "New Directions in Equity Research," *Journal of Personality and Social Psychology*, *25*(2), 151-176.

Table 1
Variation in Stakes

Treatment	Salient Stakes	Share of Stakes Claimed by First Mover: Mean	Std Dev	N	% Equal	% Self-Interested
Forsythe, Horowitz, Savin and Sefton (1994)						
1. Hypothetical $5 Dictator	$0	.617 [0236]	.184 [456]	46	.48 [23]	.13 [25]
2. Actual $5 Dictator	$5	.778 [01456]	.204 [456]	45	.18 [1456]	.36 [1456]
3. Actual $10 Dictator	$10	.767 [01456]	.179 [56]	24	.21 [156]	.21 [456]
4. Hypothetical $5 Ultimatum	$0	.562 [023]	.134 [126]	47	.43 [26]	.04 [023]
5. Actual $5 Ultimatum	$5	.545 [0123]	.104 [123]	43	.53 [23]	.00 [0123]
6. Actual $10 Ultimatum	$10	.533 [123]	.082 [1234]	24	.71 [0234]	.00 [023]
Hoffman, McCabe and Smith (1996)						
1. $10 Ultimatum	$10	.567 [0]	.076 [2]	24	.50	.00
2. $100 Ultimatum	$100	.559 [0]	.119 [1]	27	.52	.04

Note: Superscripted values indicate that the value in question differs at the five percent level of significance from its counterparts in the treatments with the superscripted numbers. For example, in the Hypothetical $5 treatment of Forsythe, et al., the superscripts [236] on the mean of .617 indicate that this mean differs from the means in treatments 2, 3 and 6 of the same study at the five percent level of significance. A superscript 0 indicates that this mean differs significantly from equal splits of .5. For tests of differences, a t-test is used for differences in means, an F-test of differences in variances for standard deviations, and a z-test of differences in proportions for percentage values, such as % Equal and % Self-Interested.

Table 2
Variation in Self-Interested Shares
Spiegel, Currie, Sonnenschein and Sen (1994)

	Salient Stakes		Self-						
	First	Second	Interest	Share of Stakes Claimed by First Mover					
Treatment	Round	Round	Share	Mean	Std Dev	N	% Equal	% Self-Interested	% Disagreements
1. Game 1	$6	$0	1.00	.686 ^{0456}	.136 ^{345}	33	.21 ^{3456}	.00 ^{2345}	.06
2. Game 2	$6	$1	.833	.686 ^{0456}	.141 ^{345}	33	.06 ^{456}	.18 ^{1346}	.00 ^{456}
3. Game 3	$6	$2	.667	.665 ^{0456}	.067 ^{126}	33	.03 ^{1456}	.61 ^{1256}	.00 ^{456}
4. Game 4	$6	$3	.500	.573 ^{012356}	.076 ^{12}	33	.45 ^{123}	.45 ^{1256}	.18 ^{23}
5. Game 5	$6	$4	.333	.490 ^{1234}	.085 ^{12}	33	.54 ^{123}	.15 ^{1346}	.18 ^{23}
6. Game 6	$6	$5	.168	.458 ^{01234}	.104 ^{3}	33	.48 ^{123}	.00 ^{2345}	.21 ^{23}

Note: Superscripted values indicate that the value in question differs at the five percent level of significance from its counterparts in the treatments with the superscripted numbers, and a superscript 0 indicates the mean differs significantly from equal splits of .5. For tests of differences, a t-test is used for differences in means, an F-test of differences in variances for standard deviations and a z-test for differences in proportions.

Table 3
Variation in Information
Share of Tickets

Treatments	Information	Value of Prizes Low	Value of Prizes High	Share of Tickets to High Prize Player Fair	Mean	Std Dev	N:Agree/Total	% Equal	% Fair	% Disagree
Roth and Malouf (1979)										
1. Games 1 & 2	Partial	$1.00	$1.00	.50	.504 [4]	.030 [4]	15/16	.80 [4]	.80 [34]	.06
2. Games 1 & 2	Full	$1.00	$1.00	.50	.491 [4]	.039 [4]	21/21	.95 [34]	.95 [34]	.00
3. Games 3 & 4	Partial	$1.25	$3.75	.25	.500 [4]	.025 [4]	16/16	.69 [24]	.00 [124]	.00
4. Games 3 & 4	Full	$1.25	$3.75	.25	.338 [0123]	.098 [123]	18/21	.17 [123]	.22 [123]	.14
Roth, Malouf and Murnighan (1981)										
1. Games 1 & 3	Low/Inter	$3.00	$9.00	.25	.509 [236]	.037 [246]	38/46	.61 [24]	.00 [6]	.17
2. Games 1 & 3	High	$3.00	$9.00	.25	.400 [0135]	.085 [1345]	24/30	.29 [1]	.08 [6]	.20
3. Game 4	Low/Inter	$3.00	$12.00	.20	.481 [126]	.043 [246]	21/23	.52	.00	.09
4. Game 4	High	$3.00	$12.00	.20	.420	.219 [1345]	11/15	.18 [1]	.00	.27
5. Game 2	Low/Inter	$2.40	$9.60	.20	.494 [26]	.052 [246]	21/23	.52	.00	.09
6. Game 2	High	$2.40	$9.60	.20	.357 [0135]	.121 [135]	12/15	.33	.17 [12]	.20
Roth and Murnighan (1982)										
1. Neither knows	Neither	$5.00	$20.00	.20	.482 [034]	.058 [234]	56/63	.59 [234]	.00 [34]	.11 [3]
2. $20 player knows	$20 player	$5.00	$20.00	.20	.461 [034]	.108 [1]	44/54	.39 [14]	.02 [34]	.19
3. $5 player knows	$5 player	$5.00	$20.00	.20	.363 [012]	.135 [1]	58/81	.28 [1]	.21 [12]	.28 [1]
4. Both know	Both	$5.00	$20.00	.20	.334 [012]	.121 [1]	51/65	.20 [12]	.16 [12]	.22

Note: Superscripted values indicate that the value in question differs at the five percent level of significance from its counterparts in the treatments with the superscripted numbers, whereas a 0 indicates that it differs significantly from equal splits. A t-test is used for differences in means, an F-test of differences in variances and a z-test for differences in proportions. The % Disagree is out of all pairs (N: Total) whereas all other figures refer only to pairs that reached an agreement (N: Agree). Two observations have been removed from the Roth, et al. (1981) data: one outlier that they also ignored and one inefficient agreement, i.e., not all tickets were allocated.

Table 4
Variation in Information
Expected Payoffs to Pair

Treatments	Information	Value of Prizes Low	Value of Prizes High	Expect Payoffs Mean to High	Expect Payoffs SD Both	N
Roth and Malouf (1979)						
1. Games 1 & 2	Partial	$1.00	$1.00	.504 [34]	.029 [34]	15
2. Games 1 & 2	Full	$1.00	$1.00	.491 [34]	.040 [34]	21
3. Games 3 & 4	Partial	$1.25	$3.75	.750 [0124]	.254 [124]	16
4. Games 3 & 4	Full	$1.25	$3.75	.594 [0123]	.138 [123]	18
Roth, Malouf and Murnighan (1981)						
1. Games 1 & 3	Low/Inter	$3.00	$9.00	.756 [02]	.259 [2]	38
2. Games 1 & 3	High	$3.00	$9.00	.659 [0135]	.181 [135]	24
3. Game 4	Low/Inter	$3.00	$12.00	.786 [02]	.291 [2]	21
4. Game 4	High	$3.00	$12.00	.709 [0]	.252	11
5. Game 2	Low/Inter	$2.40	$9.60	.794 [02]	.300 [2]	21
6. Game 2	High	$2.40	$9.60	.671 [0]	.208	12
Roth and Murnighan (1982)						
1. Neither knows	Neither	$5.00	$20.00	.784 [034]	.289 [34]	56
2. $20 player knows	$20 player	$5.00	$20.00	.762 [04]	.280 [4]	44
3. $5 player knows	$5 player	$5.00	$20.00	.664 [01]	.221 [1]	58
4. Both know	Both	$5.00	$20.00	.644 [012]	.194 [12]	51

Note: Superscripted values indicate that the value in question differs at the five percent level of significance from its counterparts in the treatments with the superscripted numbers, and a superscript 0 indicates the mean differs significantly from equal splits of .5. For tests of differences, a t-test is used for differences in means and an F-test of differences in variances for standard deviations.

Table 5
Egocentric Information Processing

	Ex Ante Stakes			Ex Post Stakes			Tests for Differences in Means	
	Mean	SD	N	Mean	SD	N	test-statistic	p-value
Babcock, Loewenstein, Issacharoff and Camerer (1995, $10 stakes)								
Fair judgments								
High stakeholder	.791	.189	47	.690	.178	47	2.67	<.01
Low stakeholder	.593	.196	47	.753	.176	47	-4.14	<.01
Difference	.198	.231	47	-.063	.248	47	5.27	<.01
t-statistic	4.96			-1.72				
p-value	<.01			.09				
Estimates of judge's award								
High stakeholder	.761	.209	47	.638	.217	47	2.81	<.01
Low stakeholder	.575	.203	47	.707	.188	47	-3.27	<.01
Difference	.186	.260	47	-.069	.286	47	4.52	<.01
t-statistic	4.37			-1.66				
p-value	<.01			.10				
Agreements								
Agreed shares	.700	.158	34	.631	.148	44	1.98	.05
Disagreement rate	.28	.44	47	.06	.25	47	2.84	<.01
Gächter and Riedl (2001, $18 stakes)								
Fair judgments								
High stakeholder	.640	.062	45	.618	.071	22	1.24	.22
Low stakeholder	.616	.068	45	.623	.066	22	-.40	.69
Difference	.024	.014	45	-.005	.021	22	5.97	<.01
t-statistic	1.75			-.24				
p-value	.08			.81				
Opening offers								
High stakeholder	.716	.054	20	.708	.105	12	.25	.81
Low stakeholder	.524	.128	25	.543	.044	10	-.65	.52
Difference	.192			.165				
t-statistic	6.78			4.95				
p-value	<.01			<.01				
Agreements								
Agreed shares	.605	.067	37	.623	.042	17	-1.20	.24
Disagreement rate	.16	.37	44	.23	.43	22	-0.69	.49

Note: The test statistic for differences in means is the t-statistic and for differences in disagreement rates the z-statistic for differences in proportions.

Table 6
Stakeholders versus Spectators
Konow (2000)

	Fair allocations		Actual allocations		Paired Difference test		
Treatment	Mean	SD	Mean	SD	t-stat	p-value	N
Discretionary differences							
1. Spectator	.491	.113	.494 [2]	.130 [2]	-0.156	.88	24
2. Stakeholder	.516	.171	.644 [1]	.235 [1]	2.959	<.01	24
Exogenous differences							
1. Spectator	.500	0	.508 [23]	.056 [2]	0.723	.48	24
2. Stakeholder	.500	0	.592 [1]	.196 [13]	2.808	<.01	36
3. Stakeholder beliefs	.500	0	.558 [1]	.083 [2]	4.205	<.01	36

Note: The t-statistics result from a two-tail test that the mean actual allocations equal the mean fair allocations. Superscripted means and standard deviations differ at the five percent level of significance from their counterparts in the treatments with the superscripted numbers.

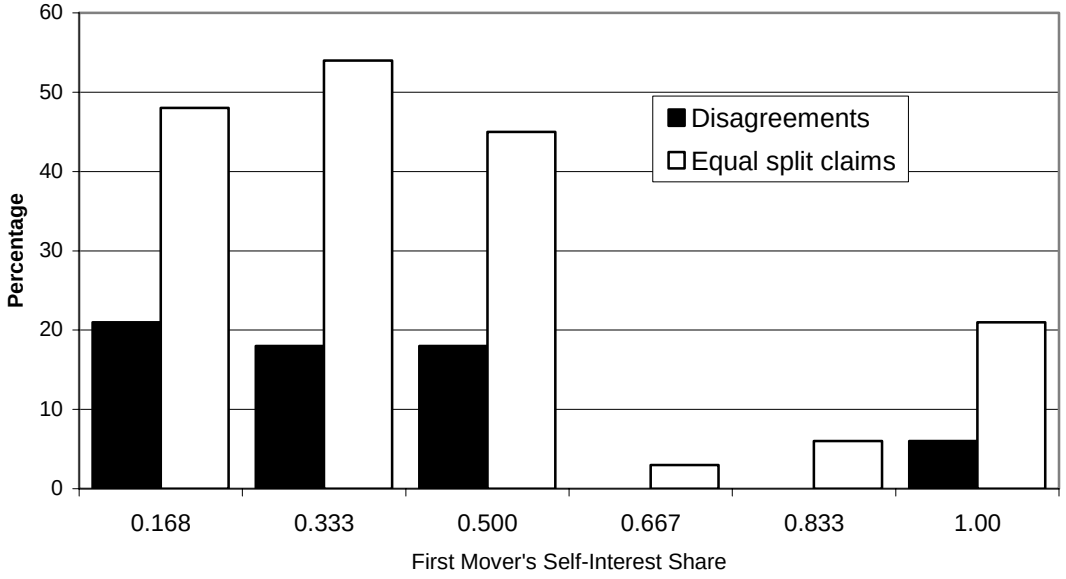

Figure 1. Equal split first mover claims and disagreements in two round bargaining experiments (Spiegel, et al., 1994).

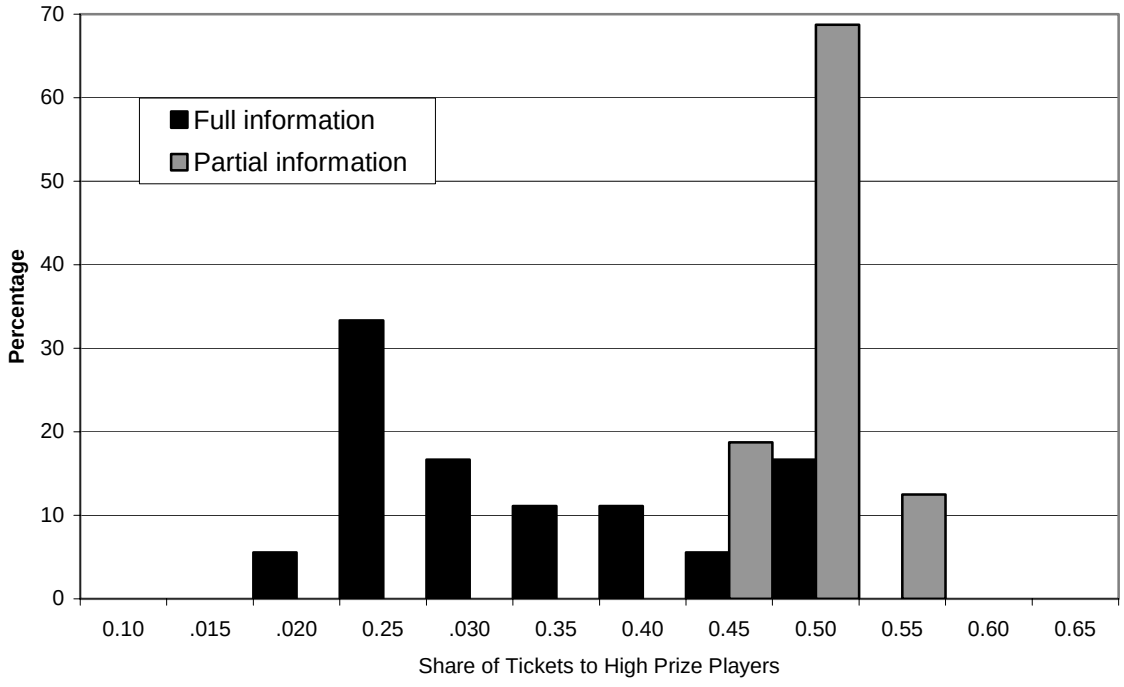

Figure 2. Agreements over lottery tickets in bargaining games (Roth and Malouf, 1979).

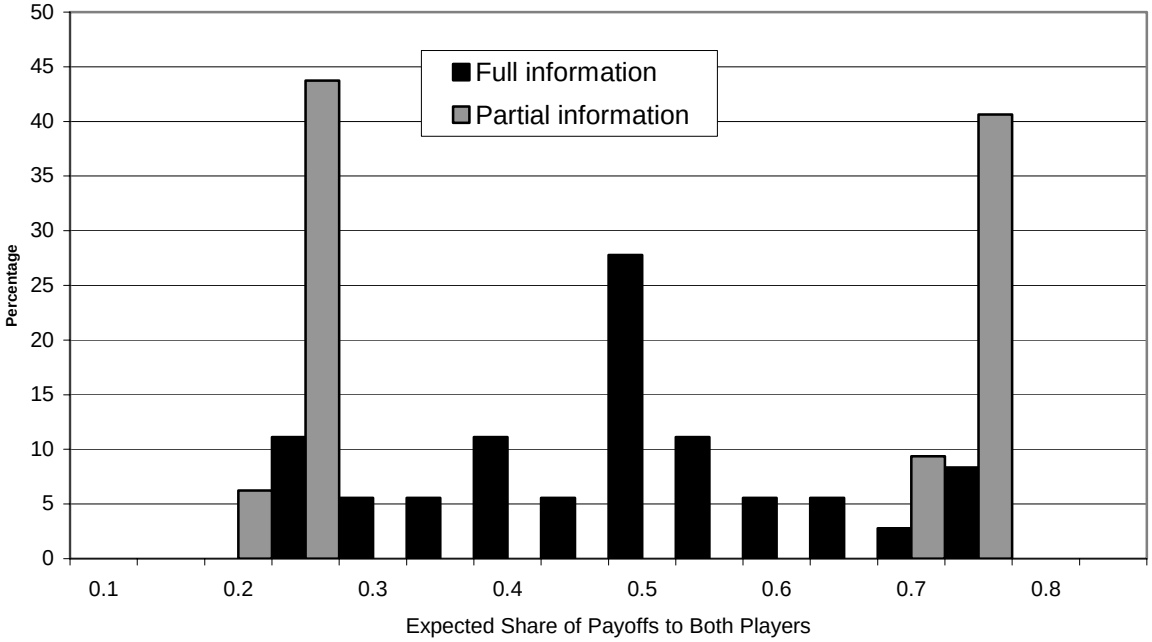

Figure 3. Expected payoffs from bargaining games (Roth and Malouf, 1979).

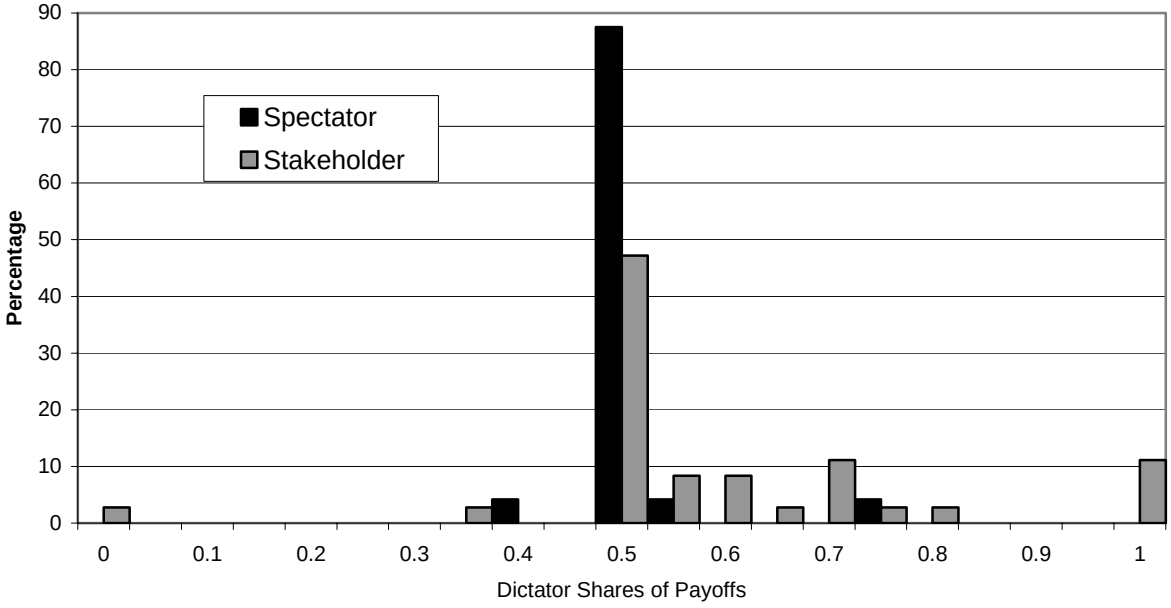

Figure 4. Dictator allocations in exogenous differences treatments (Konow, 2000).