TIME TO ABANDON THE SELF-INTERESTED UTILITY FUNCTION? - A CRITICAL NOTE ON THE THEORY OF INEQUITY AVERSION

Andreas Bergh[1]
The Ratio Institute

Prel Draft. Comments welcome!

P.O. Box 5095 | SE-102 42 Stockholm | Sweden
Phone +46-8-587 054 00 Fax +46-8-587 054 05

Paper prepared for the 2006 annual meeting of the American Public Choice Society, New Orleans, Louisiana, March 30–April 2, 2006

Category: Fairness/Justice/Equality

Abstract

The impact of the paper "A Theory of Fairness, Competition and Cooperation" by Ernst Fehr and Klaus Schmidt published in Quarterly Journal of Economics in 1999, has been tremendous. In this paper, I examine the merits of their theory of inequity aversion as a theory of fairness and as an explanation of human behavior. I claim that the theory has weak points in both areas. First, it provides no deeper understanding of why people sometimes exhibit other-regarding preferences. Second, the entirely outcome based nature of the theory means that it ignores the fundamental role of procedures, both in the theoretical literature on fairness, and in experiments regarding actual human behavior. Finally, I suggest an alternative way of understanding the puzzling behavior of humans in economic experiments, based on the potentially conflicting norms of individual property rights and social sharing.

---

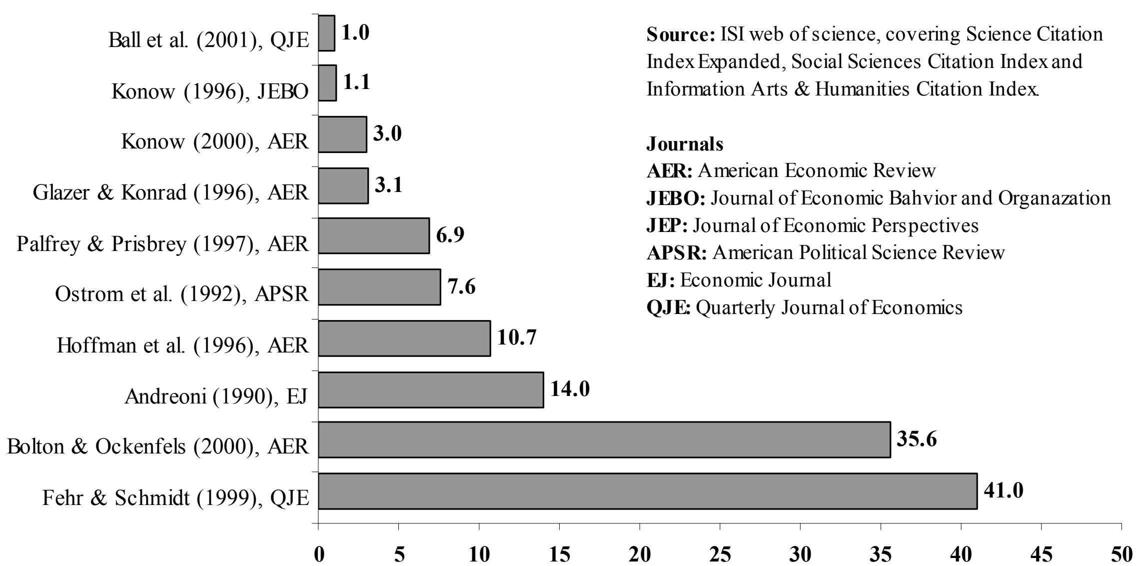
[1] Andreas.bergh@ratio.se The author would like to thank Niclas Berggren for valuable comments on earlier versions of this paper.

1. INTRODUCTION

In 1999, Quarterly Journal of Economics published an article titled "A Theory of Fairness, Competition and Cooperation" by Ernst Fehr and Klaus Schmidt, henceforth FS. The theory presented in this paper aims at explaining the high degree of other-regarding preferences as observed in several economic experiments, described in more detail below.

The impact of the FS-paper has been huge: As of October 2005 the paper has been cited an impressive 246 times in other papers.[2] This is, by a margin, more than a number of highly substantial contributions to the same research field published in top journals. Correcting for the number of years since publication, the citation impact of the FS-paper compared to some other important papers dealing with other-regarding preferences is shown in figure 1.

Average number of references per year since publication for selected journal articles



In this note, I examine the merits of inequity aversion as a theory of fairness, and as an explanation of human behavior. I argue that the merits of the theory are weak in these areas, and suggest a broader framework for understanding the type of behavior that FS aim to explain.

---

[2] According to ISI Web of science, based on Social Science Citation Index, Science Citation Index Expanded and Arts & Humanities Citation Index.

Undertaking a project like this deserves some justification. First of all, critical scrutiny of popular ideas is essential to scientific progress. Secondly, the field of behavioral and experimental economics has recently become extremely popular, and while the field has produced an impressive amount of empirical and experimental results, the supply of theoretical explanatory frameworks is rather limited. For this reason, the claim of FS to have a theory that explains human fairness behavior deserves critical scrutiny and comparisons with other theoretical approaches.

I proceed as follows. The next section briefly summarizes the most important puzzling human behavior that needs to be explained. Section three describes the FS-theory of inequity aversion, and discusses its merits as a fairness theory and as an explanation of human behavior. Section four presents a sketch of a broader framework for understanding why and when inequity aversion may occur. Section five contains some concluding remarks.

2. THE PUZZLING BEHAVIOR OF HUMANS

It has long been know that human behavior in some situations deviates substantially from the predictions of rational game theory. From experimental economics we have huge amounts of data on behavior inconsistent with short run, selfish rationality. Some important examples are given below:

- People voluntarily contribute to public goods (Marwell and Ames, 1979). Typically, free riding increases when the game is repeated – but free-riding can be mitigated by allowing for costly punishment of free riders (Fehr and Gächter, 2000).
- In the ultimatum game[3], people reject money offers from others when these offers imply too high inequality of final payoffs (Güth et al., 1982).

[3] The ultimatum game is played by a proposer and a responder. The proposer is endowed with a sum of money, and proposes a division of the sum between herself and the responder. The responder either accepts the division or rejects, in which case both receive nothing.

- In the dictator game[4] analyzed for example by Forsythe et al. (1994), people voluntarily give away money to other subjects, despite being allowed to keep the money for themselves.
- People spend resources on punishment of deviations from equal division, even when they themselves do not suffer from these deviations (Fehr and Gächter, 2002).
- People spend resources on punishment of those who violate contracts (Fehr et al., 1997) and those who overuse a common pool of resources (Ostrom et al., 1992).
- People tend to reward generous contracts by exerting more effort (Fehr et al., 1996).
- People achieve outcomes that Pareto-dominate Nash-equilibria through trust and reciprocity/trustworthiness – see Berg et al. (1995).

These findings have led to a debate on how to best explain the puzzling behavior of humans in the lab. As pointed out by FS, there are also several studies in which competitive equilibria predicted by standard rationality assumptions are confirmed in experiments. Thus, the question that needs to be answered is: Why is it that human behavior sometimes is accurately predicted by standard rationality assumptions and competitive markets, and sometimes seems to be guided by other social norms, such as fairness, reciprocity or altruism? Many of the papers mentioned in figure 1 contain important attempts to at least partially answer this question. Below is a very brief review of these papers:

- Ball et al. (2001) found that social status (achieved through assigned stars or a trivia quiz) affect behavior in a way that effect market prices, benefiting those with high status.
- Konow (1996) presents the accountability principle which "requires that a person's fair allocation (e.g., of income) vary in proportion to the relevant variables which he can influence (e.g., work effort), but not according to those which he cannot reasonably influence (e.g., a physical handicap)". (p. 13)

[4] In the dictator game, one of two players is endowed with a sum of money, and is allowed to give some, all or nothing to the other player. Then the game ends.

- Konow (2000) experimentally confirms the accountability principle, using an experiment where different outcomes may depend on either effort or luck, and also detects a self-serving bias in people's conception fairness.
- Glazer and Konrad (1996) provide a signaling/reputation explanation of charity by noting that when charitable contributions are published by size category, contributions tend to be near the minimum amount necessary to get into a category, and that hardly any giving is anonymous.
- Palfrey and Prisbrey (1997) use public good experiments and confirm weak, heterogeneous warm-glow effects (se Andreoni 1990 below) and statistically reject the presence of an altruism effect in contributions
- Ostrom et al. (1992) showed experimentally that communication, regardless of sanction possibilities, has a big positive effect on peoples' possibilities to avoid the tragedy of the commons.
- Hoffman et al. (1996) showed that other-regarding behavior in the dictator game is decreasing in the degree of anonymity.
- Andreoni (1990) notes that models with pure altruism lack predictive power and suggests instead that individuals derive utility from the act of giving ("*warm glow*").

With one possible exception (Andreoni, 1990) these papers are published in top journals. But the citation impact is still not nearly as impressive as the impact of the FS-paper. Possibly, the big citation impact is caused by the FS-paper being heavily criticized by other scholars. This seems not to be the case. On the contrary, the model has had a substantial impact and has recently been used for example to analyze economics of the welfare state (Fong, 2005), marriage and divorce (Smith, 2005), employment relations (Irlenbusch and Sliwka, 2005), tax decentralization (Güth et al., 2005), problems of optimal control (Cabrales and Charness, 2003) and the problem of collective action (Ahn et al., 2003).

This big impact and the widespread usage motivates a thorough analysis of the FS-model and its explanations of human behavior.

3. THE FS-THEORY OF INEQUITY AVERSION

The idea of inequity aversion is that people care not only about their own outcome, but also about the outcome of others. This is captured by the "Inequity Aversion Utility function". With two agents called 1 and 2, the utility of agent 1 is written as follows:

$$(1)\quad U_1(x_1, x_2) = x_1 - \alpha_1 \max\{x_2 - x_1, 0\} - \beta_1 \max\{x_1 - x_2, 0\}$$

Here, x_1 and x_2 are the monetary outcomes, α_1 measures agents 1's disutility of having less than agent 2, and β_1 measures agents 1's disutility of having more than agent 2. FS assume that $\alpha>=\beta$, so that people dislike inequity more when it is to their disadvantage.

What FS essentially do, is to show that there exist distributions of α and β values in the population such that the utility function is consistent with observed behavior in the ultimatum game, a market game with proposer competition, a market game with responder competition, and a public good game. They also state that their model explains observed behavior in gift-exchange games.

Regarding the dictator game, FS admit that their model is incompatible with data: The FS-model predicts that offers will be either 0 or exactly 0.5, which is typically not the case. They are, however, fast to suggest a solution:

"If we allow for a utility function that is concave in the amount of advantageous inequality, there is no problem in generating optimal offers that are in the interior of [0, 0.5]" (p. 848)

Note that my description of what FS do, is intentionally very conservative. Shaked (2005) has forcefully critized the rethoric in FS writings, showing that in a later paper (Fehr and Schmidt, 2003), FS falsely claim to have calibrated the model using the ultimatum game and then used the calibrated model to produce accurate predictions across many bargaining, market and co-operation games (Shaked, 2005, p. 5-6). The debate between Shaked and FS is available in volume 8 and 9 of NAJ Economics (www.najecon.org). Editors Jeff Ely and

David K. Levine note in their summary of Shaked's final comment that the debate boils down to one serious concern:

"When a researcher selects parameter values for a theoretical model consistent with data from already existing experiments, to what extent has it been shown that the model 'explains' the data?"[5]

I will continue the debate by taking on this very question: To what extent does the FS-model work as an explanation of human behavior?

3.1 Explaining human behavior

In line with Simon (1980), my basic premise here is that explaining human behavior requires general descriptions rather than casual observations. Furthermore, as argued by Bunge (1997), in order to explain and fully understand behavior, even general descriptions are insufficient: We need to understand the underlying mechanisms of behavior. Simply put, a scientific explanation of behavior needs to provide an answer to the question of why it occurs – see Woodward (2003).

So how does the contribution by FS stand up against these criteria for explanations? The bottom line of the debate between FS and Shaked, is that the contribution by FS in their QJE-paper is the invention of a utility function which for certain parameter distributions is consistent with observed behavior in a number of experiments. The experimental data were available in other papers when FS constructed their utility function.[6]

I will first discuss to what extent the FS-model works as a description of human fairness behavior, and then go on to discuss to what extent the FS-model is an explanation that gives us a deeper mechanismic understanding of human behavior.

---

[5] Quoted from www.najecon.org/naj/v9.htm.

[6] The FS-Shaked debate does not completely reveal for which experiments constistency holds. For example, Shaked raises doubts regarding the public good game where the FS-model predicts close to 100 percent freeriding, whereas data indicate a substantial amount of contributions even in the last rounds. In their reply, FS persist in their view that data are consistent with the FS-model (p. 7).

3.2 The FS-theory as a description of human fairness behavior

The situation of an economic experiment is a special one: Subjects, typically students, are recruited to participate in an experiment with real money, given to them by the experimenter. The fact that the money is given to them by the experimenter is of fundamental importance. To see this, consider what would happen if an experimenter asks experimental subjects to play ultimatum or dictator games with their own money. Granted that this has not been tested, it seems likely that very few people will agree to play these games with their own money, and that little will be transferred among those who play. A potential explanation for this, supported by evidence below, is that both persons are likely to agree on having no legitimate claims on each others money, because these money are earned.

The money people can potentially earn by participating in experiments, has a different nature: It is typically not earned, and no previously agreed upon rules exist regarding the legitimate claims to this money.

The idea implicit in the FS-model, is that peoples' conception of a fair distribution is completely consequentialistic: The utility in equation (1) depends only on outcomes and is independent of any procedures preceding the outcome.[7] As a theory of fairness, the FS-model is highly simplified. Obviously, an important merit of all economic models is indeed their simplicity: They help us focus attention on the important mechanisms of behavior. Is this the case for the FS-model? Two important strands of research suggest that the answer to this question is no: In general, fairness can not be reduced to outcomes.

First of all, the theoretical research in the normative field of distributive justice, has long ago abandoned simple outcome based models of fairness. At least since the seminal contributions by Rawls (1971) and Nozick (1974), this debate has been focused around concepts such as procedures, responsibility, opportunity and the distinction between effort and brute luck – see Roemer (1996) for an excellent overview.

---

[7] Of course, the standard self-centered utility function is also completely consequentialistic, but this is less problematic since it is rarely said to capture behavior influenced by fairness norms.

Of course, one could argue that is highly unfair to compare a utility function used by economists to highly complex theories of distributive justice within the field of political philosophy. However, there are a number of prominent example of economists who have carefully studied the field of distributive justice and managed to produce simplified economic models by focusing on the most relevant factors. One example is the theory of equality of opportunity as formulated by Roemer (1998). Another is Konow's positive theory of economic fairness, see Konow (1996, 2000). These contributions share the fact that they give a fundamental importance to procedures and context, while at the same time being simple and tractable.

Secondly, FS ignore substantial experimental evidence indicating that procedures actually do matter for fairness. For example, Hoffman et al. (1994) showed that when the role of proposer in the ultimatum game was earned, rather than being randomly assigned, proposers offered less – and respondents accepted more unequal offers. Cherry et al. (2002) conducted a dictator game with earned surplus and high anonymity, and observed that 95% of dictators keep everything.

In contrast to the work on the importance of procedures and earned entitlements, FS calibrate their model using data from experiments where money is unearned by the subjects. Thus their model applies only to the case when money appears like manna from heaven. Surely, this case is interesting – but also very special. As noted by Gächter and Riedl (2002) the importance of separating the case of earned surplus from the case of manna from heaven is also indicated by the widespread usage of different earnings procedures in experiments: Doing a concentration test (Mikula, 1972), proof reading (Frolich and Oppenheimer, 1992), cracking walnuts (Fahr and Irlenbusch, 2002), solving the Tower of Hanoi problem (Rutström and Williams, 2000), preparing letters for mailing (Konow, 2000) and solving a two-variable optimization problem (Bosman and Van Winden, 2002). When it comes to describing and explaining fairness behavior, the case when money appears like manna from heaven should reasonably be considered a special case rather than a typical one.[8]

3.3 The FS-theory as an explanation of human behavior

The FS-theory is clearly an example of an attempt at inductive explanation. FS construct their model to be compatible with observed behavior, and the model can subsequently be used to generate testable predictions regarding human behavior in other situations. We have already seen that the theory is based on observed behavior in a very special situation only, and FS provide no explanation of why their model seems to apply in the special case of unearned money, but not when money is earned through procedures. Most importantly, the FS-theory has weak explanatory value, because it provides no answers to why people exhibit inequity aversion.

The existence of other-regarding, seemingly irrational behavior, including for example altruism and prosocial cooperation, is nothing new in the social sciences. There are several attempts to explain why or under what circumstances it occurs: The accountability principle, based on equity theory (Konow, 1996, 2000), signaling (Glazer and Konrad, 1999), warm glow effects (Andreoni, 1990; Palfrey and Prisbrey, 1997), sanctions (Ostrom et al., 1992); and social distance (Hoffman et al., 1996).

The FS-model is consistent with experimental behavior under some assumptions regarding the distribution of α and β parameters in the population. This triggers a number of questions: Why are people inequity averse? Why are some people more inequity averse than others? Why the restriction that $\alpha >= \beta$? Why are α and β positively correlated?

---

[8] Note also recent evidence that both the FS-model and the model by Bolton and Ockenfels (2000) do not capture the manna from heaven-case entirely correctly when the number of participants is increased to three - see Kagel and Wolfe (2001) and Bereby-Meyer and Niederle (2005).

To the extent that these questions are adressed at all by FS, the answers given seem to be hastily produced. For example, regarding the correlation between α and β, FS answer in their reply to Shaked:[9]

> *"It seems plausible that somebody who cares a lot about inequality to his advantage also cares a lot about inequality to his disadvantage, so we assumed for simplicity that they are perfectly correlated." (p. 7)*

Furthermore, in a general journal like the *Quarterly Journal of Economics*, one might expect that proponents of a new theory of human behavior to some extent would dwell on questions regarding social efficiency and individual rationality, for example by discussing the socially optimal degree of inequity aversion, or perhaps analyze under what circumstances is it beneficial for individuals to exhibit inequity aversion. FS do not.

Under the section titled "Related approaches in the literature", FS discuss mainly the work of Rabin (1993) on fair intentions and the Bolton and Ockenfels (2000)[10] version of inequity aversion. These contributions, especially the latter, are very similar to the FS-model. Apart from this, there is no attempt to relate the FS model to a broader literature, or to compare the strengths and weaknesses of the FS explanations with those suggested by other scholars.

The sparse references to related literature is somewhat surprising. In a broader context, the behavior noted by FS can be more easily understood. Because the FS-paper only discusses experiments where surplus is generated by chance, the FS-model can be seen as a special case of Konow's (1996) positive theory of fairness. In the next section, I sketch how a broader approach, drawing on existing research, can lead to improved understanding of human behavior.

---

[9] The answer provided in the original QJE-paper (p. 864) is less informative.

[10] Which at the time was available only as a discussion paper.

3. UNDERSTANDING AND GENERALIZING INEQUITY AVERSION

FS have shown that in experimental situations with unearned money, people are prone to sharing. On the other hand, several experiments indicate strong entitlement effects of surplus earned through different procedures. Thus, human behavior is guided by at least two very different norms: Property rights and social sharing.

Both these norms may be socially efficient under some circumstances. Sharing surplus generated by chance – manna from heaven – is socially efficient because it reduces consumption variance, as has been noted and documented in several food sharing studies, for example Gurven (2004) and Kaplan and Hill (1985). It bears emphasizing that social efficiency is derived only from individual preferences: When individual utility is concave, decreasing marginal utility implies that the consumption smoothing made possible through social sharing makes everyone better off. The efficient situation need not be stable, however: For each individual there is a temptation to violate sharing norms by never sharing but still accepting gifts from others. Thus, stability requires some kind of enforcement mechanism.

When surplus is generated by individual effort, there are efficiency reasons for respecting property rights: If people there is no commonly respected norm that people are entitled to the fruits of their labor, incentives for productive efforts are weak. Furthermore, without property rights everybody must spend resources on defending their own property, which is inefficient in the sense that everybody would be better off knowing that property rights are respected. However, unless violations of the property rights are enforced, efficient outcomes are unstable because they give rise to an incentive to steal from others – see Bos and Kolmar (2003) and the references cited therein.

As already noted, the FS-model captures the effect of sharing norms when the surplus is generated by chance. Charness and Rabin (2000) provide interesting evidence for the sharing norm. Given the choice between the payoff allocations (575,575,575) and (900,300,600), where the numbers are the payoffs of players A, B and C respectively, a majority of the C-players preferred the first allocation – despite the fact that they would be better off under the second allocation. Note that the total amount is higher under the second allocation, but with a sufficiently concave utility function, total utility is higher under equal division.

For further experimental evidence of the norm of equal sharing, see Fehr and Fischbacher (2004) and Ciriolo (2006). To sum up, there is substantial experimental evidence that people do comply with the norm of sharing equally, in anonymous experiments where the surplus is generated by chance. But, as already mentioned, there are also many experiments where some kind of procedure is used to create a surplus generated by effort rather than luck. For example, Hoffman et al. (1994) showed that when a the role of proposer in the ultimatum game was earned, rather than being randomly assigned, proposers offered less – and respondents accepted more unequal offers. Thus, when the surplus to be divided is earned, behavior tilts from the sharing norm towards the norm of respecting property rights. As suggested by Cherry et al. (2002), the assets in a bargain must be legitimate to produce rational behavior.

The experiment carried out by Konow (2000) provides clear support for both sharing norms and property rights at the same time. Here, dictators were given the task of allocating a surplus between pairs who earned money by preparing envelopes for mailing. In one treatment, subjects earned differently because the amount earned per envelope was different. In another setting, the payment per envelope was equal, and the only inequality source was the amount of effort that subjects put into preparing as many envelopes as possible. When income differences were exogenous, dictators typically divided the surplus equally. When differences were caused by different effort levels, dictators typically allocated the surplus proportionally to the number envelopes prepared.[11]

Konow interprets the results as supportive of his positive theory of economic fairness – see Konow (1996). This theory "requires that a person's fair allocation (e.g., of income) vary in proportion to the relevant variables that he can influence (e.g., work effort) but not according to those that he cannot reasonably influence (e.g., a physical handicap)." (p. 13).

[11] One might of course argue that some persons are especially skilled in preparing envelopes and that therefore they benefited from this particular feature of the experiment. This can explain why there is some (but very small) degree of clustering around equal divisions also in the effort treatment.

The idea that individuals ought to be compensated for disadvantageous inequalities caused by circumstances beyond her control, but deserve to be held responsible for freely chosen actions appears in many modern theories of justice, see for example Dworkin (1981) and Roemer (1998). This is interesting, because a compensation scheme for brute luck events is essentially a social insurance, and can potentially be motivated by the insurance motive for sharing. Similarly, property rights enable society to give a person entitlement to the surplus produced as a result of her effort, thereby creating incentives for efficient use of resources. Thus, Konow's positive theory of economic fairness, as well as all similar theories of distributive justice, come very close to promoting an efficient mix of sharing norms and property rights.

While many questions are left for further research, this sketch shows that once the findings in the FS-paper are put in a broader context, we gain two important insights. The first is that the phenomenon of inequity aversion can be understood as a partial compliance with socially efficient sharing norms. The second is that is that when procedures create entitlement effects, behavior is less guided by sharing norms and more guided by property rights, which is the more efficient norm when surplus is generated by effort rather than appearing like manna from heaven.[12]

4. CONCLUSIONS

To conclude this paper, I will show two examples of how the FS-paper is cited when their model is applied by other scholars. The following passage is quoted from Irlenbusch and Sliwka (2005), discussing a typical principal-agent model where principals set wages, agents respond by choosing effort levels that feed back on the principal's payoff:

> *"In this model, we adopt the Fehr and Schmidt (1999) type of approach to explain reciprocal behavior by assuming that some agents dislike inequity. Although this approach does not model preferences for intentional reciprocal behavior directly, it*

[12] Among the questions left for further research are for example the following: What determines the relative importance of the two different norms? How do people handle potential conflicts between property rights and sharing norms? How can people solve the problem of norm enforcement to achieve stable efficient equilibria?

has been quite successful in being able to explain behavior in experiments. Inequity averse agents behave reciprocally for the following reason: suppose that an agent has received a high wage. If he does not respond with a high effort level, the income arising from the deal would be distributed unevenly (to his advantage), and his utility would be reduced as he dislikes inequity." (p. 387)

The second example comes from Güth et al. (2005), discussing the issue of tax(de)centralization:

Recent fairness research has shown that people are strongly concerned with income inequalities. In this section, we will show that a centralized tax system is "less fair" (i.e., induces more income inequalities) than a decentralized system when taxes to support local public goods are raised globally. In line with Fehr and Schmidt (1999), we capture the notion of fairness by absolute differences in payoffs.(p. 184)

The first quotation illustrates the problems when a model is said to explain a certain behavior, merely because it can be described by a utility function. The second quotation illustrates how the notion of fairness is reduced to absolute differences in outcomes in a context which is reasonably much more complex – a tax system.

At this point, I have hopefully convinced the reader that the FS-model might work as a description of human behavior in situations where a surplus is generated by chance. It does, however, not work very well to describe behavior when procedures create entitlement effects, and in any case it does not qualify as a mechanismic explanation of human behavior. My conclusion is that to foster a deeper understanding of human fairness behavior, the FS-model should be used very restrictively.

References

Ahn, T.K., Elinor Ostrom, and James M. Walker. 2003. "Heterogeneous Preferences and Collective Action." *Public Choice* 117:295-314.

Andreoni, J. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm Glow Giving." *Economic Journal* 100:464-477.

Ball, Sheryl, Catherine Eckel, Philip J Grossman, and William Zame. 2001. "Status in markets." *The Quarterly Journal of Economics* 116:161-188.

Bereby-Meyer, Y. and M. Niederle. 2005. "Fairness in bargaining." *Journal of Economic Behavior and Organization* 56:173-186.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10:122-142.

Bolton, G. and A. Ockenfels. 2000. "A Theory of Equity, Reciprocity and Competition." *American Economic Review* 90:166-193.

Bos, D and M Kolmar. 2003. "Anarchy, efficiency, and redistribution." *Journal of Public Economics* 87:2431-2457.

Bosman, Ronald and Frans Van Winden. 2002. "Emotional Hazard in a Power-to-Take Game Experiment." *Economic Journal* 112:147-169.

Bunge, Mario. 1997. "Mechanism and explanation." *Philosophy of the Social Sciences* 27:410-466.

Cabrales, Antonio and Gary Charness. 2003. "Optimal Contracts, Adverse Selection & Social Preferences: An Experiment." *Department of Economics, UCSB. Working Paper 11-03*.

Charness, Gary and Matthew Rabin. 2000. "Some Simple Tests of Preferences and a New Model."

Cherry, Todd L, Peter Frykblom, and Jason F Shogren. 2002. "Hardnose the Dictator." *American Economic Review* 92:1218-21.

Ciriolo, Emanuele. 2006. "Incorporating Fairness into the Trust Game." *Forthcoming in European Journal of Political Economy*.

Dworkin, R. 1981. "What is Equality? Part II: Equality of Resources." *Philosophy and Public Affairs* 10:283-345.

Fahr, René and Bernd Irlenbusch. 2002. "Fairness as a Constraint on Trust in Reciprocity: Earned Property Rights in a Reciprocal Exchange Experiment." *Economic Letters* 66:275-282.

Fehr, Ernst and U. Fischbacher. 2004. "Third-party punishment and social norms." *Evolution and Human Behavior* 25:63-87.

Fehr, Ernst and Simon Gächter. 2000. "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives* 14:159-181.

—. 2002. "Altruistic punishment in humans." *Nature* 415.

Fehr, Ernst, Simon Gächter, and G. Kirschsteiger. 1996. "Reciprocal Fairness and non compensating wage differentials." *Journal of Institutional and Theoretical Economics* 152:608-640.

—. 1997. "Reciprocity as a contract enforcement device: experimental evidence." *Econometrica* 65:833-860.

Fehr, Ernst and Kaus M. Schmidt. 2003. "Theories of Fairness and Reciprocity: Evidence and Economic Applications." Pp. 208-257 in *Advances in Economic Theory, Eigth world*

congress of the Econometric Society, vol. 1, edited by M. Dewatripont, L. P. Hansen, and S. Turnovski. Cambridge: Cambridge University Press.
Fong, Christina M. 2005. "Prospective Mobility, Fairness, and the Demand for Redistribution." *Mimeo, 4 January 2005*.
Forsythe, R. , J. Horowitz, N.E. Savin, and M. Sefton. 1994. "Fairness in Simple Bargaining Experiments." *Games and Economic Behavior* 6:347-369.
Frolich, Norman and Joe A. Oppenheimer. 1992. *Choosing Justice. An Experimental Approach to Ethical Theory*. Berkely: University of California Press.
Glazer, Amihai and Kai Konrad. 1996. "A signaling explanation for private charity." *American Economic Review* 86:1019-28.
Gurven, Michael. 2004. "Reciprocal altruism and food sharing decisions among Hiwi and Ache hunter–gatherers." *Behavioral Ecology and Sociobiology* 56:366-380.
Güth, Werner, Vittoria Levati, and Rupert Sausgruber. 2005. "Tax morale and (de-)centralization: An experimental study." *Public Choice* 125:171-188.
Güth, Werner, Rolf Schmittberger, and Bernd Schwarze. 1982. "An Experimental Analysis of Ultimatum Bargaining." *Journal of Economic Behavior and Organization* 3:367-388.
Gächter, Simon and Arno Riedl. 2002. "Moral Property Rights in Bargaining." *Working Paper No. 113* Institute for Empirical Research in Economics, University of Zurich.
Hoffman, E., Kevin McCabe, Keith Shachat, and Vernon Smith. 1994. "Preferences, Property Rights, and Anonymity in Bargaining Games." *Games and Economic Behavior* 7:346-380.
Hoffman, Elizabeth, Kevin McCabe, and Vernon Smith. 1996. "Social Distance and Other-Regarding Behavior in Dictator Games." *American Economic Review* 86:653-660.
Irlenbusch, B. and D. Sliwka. 2005. "Transparency and reciprocal behavior in employment relations." *Journal of Economic Behavior and Organization* 56:383-403.
Kagel, John H. and Katherine Willey Wolfe. 2001. "Tests of Fairness Models Based on Equity Considerations in a Three-Person Ultimatum Game." *Experimental Economics Year* 4:203-219.
Kaplan, H and K Hill. 1985. "Food sharing among Ache foragers: tests of explanatory hypotheses." *Current Anthropology* 26:223-245.
Konow, James. 1996. "A Positive Theory of Economic Fairness." *Journal of Economic Behavior and Organization* 31:13-35.
—. 2000. "Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions." *American Economic Review* 90:1072-1091.
Marwell, G. and R. E. Ames. 1979. "Experiments on the Provision of Public Goods. I. Resources, Interest, Group Size, and the Free-Rider Problem." *American Journal of Sociology* 84:1335-1360.
Mikula, Gerold. 1972. "Gewinnaufteilungsverhalten in Dyaden Bei Variiertem Leistungsverhältnis." *Zeitschrift für sozialpsychologie* 3:126-133.
Nozick, R. 1974. *Anarchy, State and Utopia*. New York: Basic Books.
Ostrom, Elinor, James Walker, and Roy Gardner. 1992. "Convenants with and without a Sword: Self-Governance is Possible." *American Political Science Review* 86:404-417.
Palfrey, Thomas R. and Jeffrey E. Prisbrey. 1997. "Anomalous Behavior in Public Goods Experiments: How Much and Why?" *American Economic Review* 87:829-846.
Rabin, M. 1993. "Incorporating Fairness Into Game Theory and Economics." *American Economic Review* 83:1281-1302.
Rawls, J. 1971. *A Theory of Justice*. Massachusetts: Harvard University Press.
Roemer, John E. 1996. *Theories of Distributive Justice*. USA: Harvard University Press.

—. 1998. *Equality of Opportunity*. Cambridge: Harvard University Press.
Rutström, Elisabeth E. and Melonie B. Williams. 2000. "Entitlement and Fairness: An Experimental Study of Distributive Preferences." *Journal of Economic Behavior and Organization* 43:75-89.
Shaked, Avner. 2005. "The Rhetoric of Inequity Aversion." *NAJ Economics* 8.
Simon, Herbert A. 1980. "The Behavioral and Social Sciences." *Science* 209:72-78.
Smith, I. 2005. "Fairness and inefficient divorce." *Economics Letters* 88:159-163.
Woodward, James. 2003. "Scientific Explanation." in *The Stanford Encyclopedia of Philosophy (Summer 2003 Edition),*, edited by E. N. Zalta.