# The Rhetoric of Inequity Aversion*

Avner Shaked†
Economics Department, Bonn University
Adenauerallee 24, 53113 Bonn
Germany

This version: 8 March 2005

**Abstract**

We read two papers on Inequity Aversion by E. Fehr & K. Schmidt. We examine the structure and quality of their arguments and the use they make of rhetoric in their papers.

*I am grateful to E. Fehr & K. Schmidt for pointing out two imprecisions in the previous version of this pamphlet, I have corrected them in this version. These were partly due to a misprint in their original paper. The main points and issues raised in this pamphlet are unaffected by these changes.
The complete exchnage between Fehr-Schmidt and Shaked can be found in: http://www.wiwi.uni-bonn.de/shaked/rhetoric/

†I wish to thank 10 friends who read an earlier version of this pamphlet, and who helped me focus my arguments and moderate my excessively outraged tone.

# The Rhetoric of Inequity Aversion

## 1 Introduction

The idea of Inequity Aversion, introduced by E. Fehr and K. Schmidt in a QJE paper, has from its inception, struck a chord in the heart of Economics. It has been particularly attractive to Experimental Economists to whom it was presumably directed. Two years after the first paper introducing the idea was written, a year after its publication, the authors have been invited to present a survey of their and related theories in the distinguished 8th World Congress of the Econometric Society. The two papers have soon become very popular, many experimentalists applying the new concept and many (well over 500, according to Scholar.Google) citing these papers. The attraction of the theory is, no doubt, due to its simplicity and its ease of use, but partly also due to Fehr and Schmidt's claim of having tested and confirmed the theory using extensive experimental data.

For my shame, I have not read this literature until recently. When I finally came to read the two papers, I was introduced to a subculture hitherto unknown to me, a subculture that apparently coexists parallel to main-line economics and in which different rules of logic and different laws of proof apply.

In this sphere it seems to be permitted to place crucial information in appendices and footnotes, to treat data in a casual and nonchalant way, and it is allowed to inflate results when citing them in subsequent papers.

All these transgressions can be found in the two papers by Fehr and Schmidt who at the same time praise the sophisticated scientific standards of Experimental Economics.

After introducing the idea of Inequity Aversion in their QJE paper, Fehr & Schmidt present two important results whose aim is to test and confirm their theory.

The first result in the QJE paper is an important test of their theory. Fehr and Schmidt calibrate their model using the data from Ultimatum Games. Using this calibration they explain the experimental observations in four other games.

The second confirmation of the theory is their claim that the theory of inequity aversion has no quarrel with V. Smith's established and robust experimental observations of Competitive Markets. Smith and others have shown that experimental markets converge to the competitive equilibrium. Fehr and Schmidt aim to show that the fairness they add to the players' preferences does not contradict these results. They claim that even in populations with high degree of inequity aversion the equilibrium is the competitive one (or close to it).

I will demonstrate that Fehr & Schmidt have failed to show these results. I will argue that by obscuring fundamental methodological questions, by treating data in a casual way and by inflating their own results, they have wrapped their statements and arguments with so much confusion and fog that it is easy for the reader to believe that Fehr and Schmidt established their important results, when, in fact, they have not. I will show that Fehr & Schmidt continue to exaggerate the extent of these results when citing them in recent papers (2003-2004).

It is not for me to say why Fehr and Schmidt resorted to such techniques, but whatever the reason may be, the effect of these rhetorical devices is to augment and boost the papers' message.

These are grave charges that ought to be, and will be, meticulously substantiated in this pamphlet. For this purpose I will quote extensively from the two papers. The first paper (QJE) introduces the theory of inequity aversion and contains the main results and computations. The second, the invited paper was written a year later and is a review of the related literature, as such, it juxtaposes the theory of Inequity Aversion with other related theories. This has allowed the authors to reflect upon their concepts and clarify their views. There are no new results concerning Inequity Aversion in the Invited Paper, but the aims and the results of the QJE paper are more clearly expressed in it.

It is NOT my aim to argue with the theory of Inequity Aversion in this pamphlet. I do not intend to examine its general methodology or how useful the theory may be, neither confirm it, nor prove it false. My aim is to examine the rhetorical devices used by the authors, to look at the structure and quality of their arguments, their logic, their choice of words and the way they use them, and to examine how the description of their results evolves over time. In a way, this pamphlet is a literary study of their papers, which, given the subject matter, assumes some knowledge of technical and economic concepts.

The two papers of Fehr and Schmidt are:

1. Fehr, Ernst and Klaus M. Schmidt, 1999. "A Theory of Fairness, Competition and Co-operation." *Quarterly Journal of Economics, 114, 817-868.*

2. Fehr, Ernst and Klaus M. Schmidt, 2003. "Theories of Fairness and Reciprocity: Evidence and Economic Applications", in: M. Dewatripont, L.P. Hansen, S. Turnovski, *Advances in Economic Theory, Eighth World Congress of the Econometric Society,* Vol. 1, Cambridge: Cambridge University Press, 208-257.

I refer to the first paper as **QJE,** to the second as **INV**, and to the two authors Fehr & Schmidt as **F&S**.

## 2 The Theory of Inequity Aversion

The starting point of the papers is the empirical and experimental evidence accumulated over the last decades, indicating that in many one-shot games people do not behave according to the traditional theory that assumes individuals to be concerned only with their own material payoffs. There is experimental evidence that in some (one-shot) situations individuals behave selfishly and in others they are willing to reward or punish others at a cost to themselves. Fehr & Schmidt propose to change the basic assumption of selfishness by introducing a utility function which will permit equity to enter the individuals' preferences. They suggest a utility function that depends on the material payoffs of ALL individuals - the distribution of payoffs in the population. This is the only deviation from the traditional theory. F&S continue to assume that, given their utility functions, the players are fully rational. Fehr & Schmidt's aim is to find a simple model that will explain this seemingly contradictory altruistic, spiteful and selfish behavior.

Before introducing their theory, F&S embark on a short debate with doubting Economists. *"Why are economists so reluctant to give up the self-interest hypothesis?"* they ask, and provide two answers. The first is that self-interest proved very successful in predicting behavior in some situations, the second is a methodological objection to change the basic assumption since anything can be explained by choosing the 'right' preferences. F&S' response to the last objection is that it made sense in the past, but now that experimental economics has developed *'sophisticated tools to examine the nature of preferences in a scientifically rigorous way'*, this is no longer true (INV p.209). The existence of these tools is meant to remove all doubts from the remaining sceptics about using the proposed utility function of F&S. In the following pages we will carefully read F&S' papers and we will get a taste of these sophisticated and rigorous scientific tools.

**The Inequity Aversion Utility function**

F&S propose the following utility function as representing their Inequity Aversion for a population of $n$ individuals:

$$U_i(x_1, x_2, ...x_n) = x_i - \frac{\alpha_i}{n-1}\sum_{j\neq i}\max\{x_j - x_i, 0\} - \frac{\beta_i}{n-1}\sum_{j\neq i}\max\{x_i - x_j, 0\},$$

$$\text{where} \quad 0 < \beta_i < 1, \quad \beta_i < \alpha_i$$

For 2 individuals, this becomes:

$$U_1(x_1, x_2) = x_1 - [\alpha_1 \max\{x_2 - x_1, 0\} + \beta_1 \max\{x_1 - x_2, 0\}].$$

The parameter $\alpha$ measures the envy of being poorer than another individual while the parameter $\beta$ measures the discomfort of being better off. The utility function is normalized by the factor $n-1$, where $n$ is the size of the population. The utility function allows an individual to behave altruistically or spitefully depending on the distribution of payoffs.

# 3 Prediction Across Games

In section V of QJE (Prediction Across Games) the authors *"examine whether the distribution of parameters that is consistent with experimental observations in the ultimatum game is consistent with the experimental evidence from the other games."* (QJE p.843) In fact, they do more: they derive a distribution of $\alpha, \beta's$ from the experimental data on Ultimatum Games and use it to explain the experimental data of 4 other games: Ultimatum Games with Proposer and Responder Competition and Public Good Games With and Without Punishments.

F&S do not tell us anything about the methodology of this computation. In the above quotation they talk about '*THE distribution of parameters that is consistent with the experimental observation in the ultimatum game'.* What do they intend to do if they find many such distributions and not all of them consistent with the evidence from the games they wish to explain? Will they be content to find a single distribution that explains the other games? How will they go about choosing this distribution? Do they have a method to choose such a distribution, a method that will avoid the obvious pitfalls of being influenced by the data of the games they wish to explain? If, indeed, it is their aim to find a single distribution, why give Ultimatum Games a special status? Simply write down a distribution and show that it explains all games, including the UG.

These are such obvious questions to ask, that it is hard to believe that F&S have not considered them. But they do not mention these questions in their papers, and if they do have any answers they do not share them with the reader.

Before we actually look at their computation, it is instructive to see how they describe this computation in their QJE paper and a year later in INV (the emphases are mine).

> In QJE: *"The objective is rather to offer a* ***first test*** *for whether there is a chance that our theory is consistent with the* quantitative *evidence from* ***different games****. Admittedly, this test is rather* ***crude****."* (QJE, p.843 top), and again: *"Clearly, the above* ***computations*** *provide only* ***rough*** *evidence in favor of our model."* (QJE p. 846)
>
> In INV: *"Using the data that is available from many experiments on the ultimatum game, Fehr and Schmidt* ***calibrate*** *the distribution of $\alpha$ and $\beta$ in the population. Keeping this distribution constant,*

*they show that their model yields quantitatively* ***accurate*** *predictions across* ***many*** *bargaining, market and co-operation games."*(INV p.222)[1]

The descriptions *'crude'*,*'rough'* and *'first test'* (in QJE) are replaced by *'accurate'*(in INV), the mere *'computation'* is upgraded to a *'calibration',* and *'different games'* becomes *'many games'.*

It must be, that during the year before writing the second paper, F&S have critically reassessed their scientific standards and found them exacting and far too demanding. They realized that they have been too severe to themselves and that the calculation they erroneously called crude, is in fact accurate. In addition, they re-read the extensive literature on calibration (the term calibration does not even appear in QJE) and decided that what they called a computation deserves the title calibration.

F&S did not tell us why they thought the computations were crude, nor do they tell us what made them change their mind. Although it would have been appropriate, indeed essential, to inform us why they think that the computations should no longer be considered crude.

It is unfair to criticize a crude and rough calculation which is meant to be only a first test, but now that F&S see this calculation as a calibration which yields accurate predictions, we may be justified in applying stringent measures to test their claims. A careful study of their computation will show that their first description of it is by far the more suitable one.

## 3.1 The Impossibility of a Fine Calibration.

Recall that the distribution of the parameters in the Inequity Aversion model is a *joint distribution* of $\alpha's$ and $\beta's$. Each individual in the population has both $\alpha$ and $\beta$. This joint distribution cannot be deduced from the data on Ultimatum games. In the ultimatum game a responder reacting to a given offer will practice either his envy ($\alpha$) or his discomfort at being better off ($\beta$) but never both simultaneously. According to the theory, a proposer will never offer more than 1/2 of the surplus and so he practices only his $\beta$. Unless we are informed how the same individual played in both roles, the most one can find using the data are the separate derived distributions of $\alpha$ and $\beta$. The data F&S present is not detailed at the individual level, so a joint distribution cannot be deduced from this data.

But even the derived distributions of $\alpha, \beta$ cannot be finely tuned. When inequity aversion theory is applied to the UG it predicts only coarse intervals for the values of $\beta$. This can be easily confirmed by a quick look at equation 5 of proposition 1 (QJE p. 827)[2]. For example, all that the theory tells us about those proposers who made an offer of 0.5 is that their $\beta$ is $\geq 0.5$.

It is therefore impossible to *fully* calibrate the inequity aversion model with the UG data. Unfortunately, F&S need the joint distribution for explaining

[1] The reference is to the QJE paper. There is no other reference to Fehr & Schmidt in INV.

[2] Equation 5 is reproduced on p.22, in the appendix of this pamphlet.

the behavior in the public good game with punishment, and the exact values of $\beta$ prove to be crucial for explaining other games. F&S do not address these difficulties. We shall see later what F&S do when they are forced to face these problems.

### 3.2 The Data on Ultimatum Games and the Calibration.

The data on Ultimatum Games that F&S use for the calibration is to be found in Table I of QJE and in Roth [1995]. Table I (QJE p.827)[3] lists the offers made by the proposers in ultimatum games. Proposers do not offer more than 1/2 of the surplus, so this information may be used to find the distribution of their $\beta's$ (the discomfort of being richer). A responder faced with an offer of less than $\frac{1}{2}$ of the surplus, uses his $\alpha$ to accept or reject the offer. Thus, to find the distribution of $\alpha's$ one needs to have detailed information on offer rejection rates by responders. This information is not in Table I, it is presumably to be found in Roth [1995]. The 1995 paper by Roth is a 95 (!) page long article in the Handbook of Experimental Economics. F&S do not tell us where exactly in these 95 pages the data can be found. The article has some scattered data on rejection rates in one or two experiments, but I could not find the data in the format F&S use. Transparency does not seem to be part of the rigor that scientific methods require. Could it be that F&S distilled the data from the 16 diagrams on pages 285-287 of Roth [1995], or perhaps from Figure 4.6 on p.278? Could it be that all the data they allude to refers only to the UG experiments of Roth et al. [1991], i.e. a mere 13% of the observations in Table I ?

In fact, it matters very little what exactly the data is, since all we are told about it is included in the following 3 sentences, and the wording of these sentences makes it clear that the choice of the parameters is very casual:

> "*Thus, we* ***(conservatively)*** *assume that* 10% *of the subjects have* $\alpha = 4$ " and "*Another, typically much larger fraction of the population insists on getting at least one-third of the surplus, which implies a value of* $\alpha$ *which is equal to one. These are* ***at least*** 30 *percent of the population.*" and "*Another,* ***say,*** 30 *percent of the subjects insist on getting at least one quarter, which implies* $\alpha = 0.5$" (QJE pp. $843 - 844$).

Clearly, the data does not pin down the distribution. There are many degrees of freedom left, there are many distributions that will fit this data. The theory assumes that proposers are familiar with the distribution of inequity aversion among the responders ($\alpha$). It postulates that they make their offers given their own inequity aversion ($\beta$) and that they optimize against the responders $\alpha$ distribution. Thus, the intervals of $\beta's$ that F&S suggest in equation 14 (QJE p.844) are a function of the coarsely chosen $\alpha$ distribution, these intervals are therefore subject to large variations.

[3] Table I is reproduced on p.20, in the appendix of this pamphlet.

To complete the distribution of the parameter $\beta$, F&S provide this information:

> *"If we look at the actual offers made in the ultimatum game, there are roughly* 40 *percent of the subjects who suggest an equal split. Another* 30 *percent offer* $s \in [0.4, 0.5)$*, while* 30 *percent offer less than* 0.4*. There are hardly any offers below* 0.25*. "* (QJE 844)

This information is not available in Table I, and F&S do not inform us where they got this data from. Table I only tells us that about 71% made offers between $[0.4, 0.5]$ without further details. Moreover, in Table I about 4% made offers below 0.2.

But even if we accept F&S' proposed $\alpha, \beta$ distribution, equation 14 determines only coarse intervals of $\beta' s$: all we know about those proposers who offered 0.5 is that their $\beta$ is greater or equals half $\left(\beta \geq \frac{1}{2}\right)$. For these 40% F&S have arbitrarily chosen the value $\beta = 0.6$. We will see that this value plays an important role in what follows.

The calibrations of the (separate) distributions of $\alpha, \beta' s$ suggested by F&S are given in Table III[4] (QJE p.844).

F&S use this calibration to explain the experimental behavior in 4 different games: A market game with proposers competition, a market game with responders competition , a public good game without punishment and a public good game with punishment

Let us now see how F&S explain these experiments.

### 3.3 Explaining the Market Game with Proposers' Competition.

The first experiment to be explained is a market game with proposers' competition, by Roth, Prasnikar, Okuno Fujiwara and Zamir [1991]. The game is an ultimatum game with added competition among the proposers and a crucial restriction of the single responder's actions. In this game a number of proposers simultaneously make offers to a single responder who is restricted to accept or reject the *highest* offer. If he accepts the offer, one of the proposers of that offer is chosen at random and they split the surplus, while all others get 0 payoff (QJE p.829). In the experiment by Roth et al. the outcome converged to the competitive equilibrium (in which the single responder gets all the surplus). F&S prove in Proposition 2 (QJE p.830) that for any degree of inequity aversion of the players, the only equilibrium of the game is the competitive one in which the responder receives the total surplus. There is no need to restrict the distribution of inequity aversion parameters to the proposed calibration, any distribution of parameters and preferences yields the competitive outcome[5]. In

[4] Table III is reproduced on p.23, in the appendix of this pamphlet.

[5] In fact, this result has little to do with inequity aversion. The result holds as long as the utility function satisfies the following property: For any given distribution of material payoffs, a player prefers to be in a position which gives him more material goods, rather than less.

section 4 we discuss the relevance of this particular game to what F&S wish to show.

### 3.4 Explaining the Market Game with Responders' Competition.

The situation is slightly different in the market game with responders' competition by Güth, Marchand, and Rulliere [1997]. Here, a single proposer makes an offer to a group of responders who simultaneously accept or reject it. Among those who accepted the offer one responder is chosen at random and the pair splits the surplus while all others receive 0 payoff. In a competitive outcome all the surplus is given to the single proposer. The experiment was conducted in groups of 6, a single proposer and 5 responders. F&S report that in the experiment conducted by Güth et al., about 80% of the responders accepted an offer of less than 0.02 of the surplus, i.e. the outcome was close to the competitive one (QJE p.832).

F&S prove in Proposition 3 (QJE p.832) that in order to guarantee an outcome close to the competitive equilibrium *two* individuals need to be sufficiently selfish: the proposer AND the least inequity averse responder. The proposer's inequity aversion parameter $\beta$ should satisfy $\beta < \frac{n-1}{n}$, where $n$ is the size of the population.

F&S learn from Proposition 3 that *"if there is at least one responder who does not care about disadvantageous inequality (i.e., $\alpha_i = 0$), then there is a unique equilibrium outcome with $\bar{s} = 0$."* (QJE p.845). F&S calculate the probability that at least one selfish responder (with $\alpha = 0$) is present in a group of 5 responders. According to the calibration this is $1 - 0.7^5 = 83\%$ , a figure which is roughly compatible with the 80% groups in the experiment who played very close to the competitive equilibrium.

The additional, totally ignored, condition for the selfishness of the proposer is $\beta < \frac{n-1}{n} = \frac{5}{6} = 0.8\bar{3}$. Miraculously, the highest $\beta$ value F&S chose in their distribution is $\mathbf{0.6} < 0.8\bar{3}$, (recall that any value $\geq \frac{1}{2}$ could have been chosen). That is, ALL proposers are made to be sufficiently selfish. F&S do not even mention in their calculations that it is their choice of the value 0.6 that validates their calculation. The significance of the choice of 0.6 is completely hidden from the reader.

However, choosing the highest $\beta$ as 0.84 (rather than 0.6) yields a probability of $(1 - 0.4)\left(1 - 0.7^5\right) = 0.499\,16$ for the outcome to be competitive, i.e. 50%. This is considerably less compatible with 80%. The casual choice of values for the proposed distribution proved to be very successful.

In fact, this miracle should be more stringently interpreted. The calibration of the inequity aversion model using the UG data fails to explain the experiment by Güth et al., there is a large set of $\beta$ distributions, all consistent with the UG data that are incompatible with the behavior in the Güth et al. experiment. By picking the highest value of $\beta$ to be $< \frac{5}{6} = 0.8\bar{3}$, F&S have used the data in Güth et al. to further calibrate their model. Thus the Güth et al experiment has been instrumental in the calibration and can no longer be explained by it.

### 3.5 Explaining Public Good Games

F&S analyze two types of Public Good Games: with and without punishments. A public good game with no punishment is a one stage game in which all players have income $y$. Each player may give up an amount $g$ of his income which becomes a quantity $ag$ ($a < 1$) of a public good enjoyed by all. Efficiency requires that all players contribute their entire income $y$, but the equilibrium of the traditional model with selfish preferences has all players contribute $g = 0$. In experiments, about 73% of the players made no contributions (QJE p.837).

The second game is a public good game with punishment, it was introduced in an experiment by Fehr and Gächter [2000] (in QJE this experiment is referred to as Fehr and Gächter [1996], but it was published in the meantime). The game has two stages and in the first stage it is like the public good game with no punishment in which players may contribute to a public good. In the second stage players may punish others at a cost to themselves. The important experimental observation is that about 80% of the players contribute their *entire* income (QJE p.838). Thus, the threat of punishment induces efficiency.

In explaining the two games the Inequity Aversion theory faces the challenge of simultaneously explaining free riding in one set-up and cooperation in another.

#### 3.5.1 Explaining Public Good Games without Punishment

The analysis of this game shows that certain individuals will never contribute to the public good and that the others may contribute if their number is sufficiently large. The details are in Proposition 4 (QJE p.839).

To explain the data of this game, F&S consider the case $k = 1$, a single sufficiently selfish individual. According to proposition $4(b)$[6], if $a(n-1)/2 \leq k = 1$, a single person with a sufficiently low $\beta$ suffices to enforce an equilibrium in which no one contributes. If there is a high probability that such an individual is present in each group, and all players know it, then they will all contribute $g = 0$ (this probability can be calculated by using the calibrated distribution of $\beta' s$).

The data on the experiments of Public Good Games with no Punishment is presented in Table II (QJE p. 838)[7].

F&S take the experiments by Croson and Keser & van Winden as *'typical treatments'* with $n = 4, a = 0.5$. For these experiments, with altogether 304 subjects, they calculate that the calibrated theory predicts about 97.44% of free riding. We are not being informed that for the three Andreoni experiments with 230 subjects the theory predicts 98.97% , $(= 100 - 0.4^{5})$of free riding, and for the rest of the experiments in Table II, with altogether 404 subjects, the calibrated theory predicts 100% of free riding. However, Table II reports that only 73% of the subjects do not contribute. This is a gap of over 33% $\left(= \frac{97.44-73}{73}\right)$ of the observations. How do F&S account for this discrepancy? They refer to

[6] Proposition 4, in QJE, contains a misprint: The condition on $k$ in $4(b)$ should have a **weak** inequality.

[7] Table II is reproduced on p.21, in the appendix of this pamphlet.

it in footnote 21 (QJE p.845). F&S inform us that in the final round of the game there is a significant fraction of subjects who contribute very little, those should be added to the 73%. Although this is the second time that F&S refer to this non-negligible fraction of subjects (see QJE p.837), they do not provide the data. F&S do not tell us why this data was not included in Table II - the sole purpose of Table II is to provide the data for explaining the behavior in public good games, if without this detailed data the experiment cannot be explained, why isn't the data there[8] ?

To be fair to F&S, partial information (relating only to the Fehr & Gächter [2000] experiment) can be found in the original figure of Fehr & Gächter reproduced on p.839 of QJE[9]. With the help of a magnifying glass one can attempt to fish the missing information from the gray puddles of the figure (these refer to the treatment without punishments): My imprecise estimate is that even if one takes all the contributions up to 40% of the endowment, the percentage of free riding is well below 70%.

It seems that this additional data does not really fill the gap, because F&S tell us (in the footnote) that they leave it for future research to explain this 33% of noise with the help of theories of *fundamental randomness of human choice*. In the meantime they conclude (in the text) that *"it seems fair to say that our model is consistent with the bulk of individual choices in this game"* (QJE p.845).

In all fairness, it seems right to say that until they provide data that substantially closes the gap, or until they do this promised research on fundamental randomness of human behavior (predicting 33% noise), it cannot be said that their model explains the data of this experiment. It is also fair to say that once a theory that predicts so much noise is formulated, few things will be left unexplained.

### 3.5.2 Explaining Public Good Games with Punishment

F&S profess that from the perspective of their theory this experiment is the most interesting one (QJE p.845), so it must be very important for them to explain it.

F&S Prove in Proposition 5 (QJE p.841), that if there is a group of *"conditionally cooperative enforcers"*, players with sufficiently high $\alpha' s$ and $\beta' s$, while all the other players are completely selfish $(\alpha = \beta = 0)$, then there exist equilibria in which *all* players contribute the same amount $g \in [0, y]$ in the first stage of the game. To choose among these equilibria F&S apply a *'reasonable refinement argument'*, they claim that the efficient equilibrium with $g = y$ is a *'natural focal point'* (QJE p.842).

[8] E. Fehr, in two papers with other authors, refers to Table II in QJE as a meta-study of public good games with no punishment. He attributes to this study more information than can be found in QJE (contributions in earlier stages of the game). He also mentions that in the last stage 'many' (no figure is given) players contributed close to zero.
See E. Fehr, H. Gintis, S. Bowles & R. Boyd: "Explaining Altruistic Behavior in Humans", *Evolution and Human Behavior*, 24 (2003), p. 160, and Fehr & Gächter [2000] p.983

[9] This figure is reproduced on p.22, in the appendix of this pamphlet.

The calculations, which according to F&S are somewhat cumbersome, are put in an appendix. It is in the appendix (QJE p.864) that we first hear of the need to know the correlation of $\alpha$ and $\beta$ in order to explain this game. It is the first time that the term 'correlation' appears in QJE in conjunction with the distribution of $\alpha$,and $\beta$.

In the appendix we are told that F&S *"mentioned already that the empirical evidence suggests that these parameters are positively correlated"* (QJE p.864). I could not find any sentence in which the correlation of the parameters was mentioned, nor could I find any other experiment from which F&S could derive this correlation. Perhaps they mean the following sentence: *"Fehr and Gächter report that the vast majority of punishments is imposed by cooperators on the defectors and that lower contribution levels are associated with higher received punishments."* (QJE p.839). This sentence informs us about some empirical findings, not about a correlation of parameters in a theory. Leaping from the empirical evidence to a property of parameters in a theory is known as calibration. F&S use this observation and assume a perfect correlation between the parameters. They ignore the fact that by using the joint distribution of $\alpha, \beta$ which was derived from the data of the Fehr-Gächter experiment they are prohibited from explaining this experiment.

But there is another morsel of calibration hidden in the calculations. Here, as in the market experiment with responders' competition, the casual choice of the maximal value of $\beta$ as 0.6 proves to be very crucial. The UG data allowed any value $\beta \geq \frac{1}{2}$. To explain the Fehr-Gächter experiment, Proposition 5 (QJE p. 841) requires the existence of *"conditionally cooperative enforcers"* with $\beta \geq 1 - a = 0.6$, in the Fehr-Gächter experiment $a = 0.4$. If the maximal value of $\beta$ is taken to be less than 0.6, the theory *totally* fails to explain this experiment. Yet again, the choice of a particular value as 0.6 and not, say, 0.59 means that F&S have used this experiment to calibrate their model. and it can no longer be explained by it.

There are other questions that the 'explanation' of this game leaves open. Although these questions are, strictly speaking, superfluous in view of the above arguments, I'll briefly mention them for completeness sake.

- F&S claim that $g = y$ is "a natural focal point" because it is efficient. Is it more natural than $g = 0$ or indeed $g = y/2$ ? These are also symmetric focal points, albeit inefficient ones. F&S do not support their choice of a particular focal point by any experimental evidence. I believe there is a debate in the experimental community concerning efficiency and inequity aversion and that it has not been settled yet.

- F&S quote and apply Proposition 5 incorrectly: *"For example, if all four players believe that there is at least one player with* $\alpha_i \geq 1.5$ *and* $\beta_i \geq 0.6$*, there is an equilibrium in which all four players contribute the maximum amount." (*QJE p.846). Proposition 5 postulates that the members of a group are either 'conditionally cooperative enforcers' or completely selfish

$(\alpha = \beta = 0)$. The condition that the other players are completely selfish as required by the proposition seems to be completely missing from the calculations.

### 3.6 Dictator Games and Gift Exchange Games

These two games seem to be in a category and a section (Section VI, QJE p.847) of their own. F&S provide intuition why the theory of inequity aversion can explain the behavior in experiments of these games, but they do not attempt to explain these games using the calibration. For dictator games F&S reason that they can be better explained by a non linear version of their model.

### 3.7 A summary of the Calibration

The data of the ultimatum game cannot be used to find the *joint* distribution of $\alpha, \beta$. For the separate distributions of $\alpha, \beta$, the theory (when applied to the UG) yields very crudely defined intervals of the parameters.

- The calibration failed to explain the Market Game with Responders' competition and the Public Good Game with Punishment. F&S have used the data of these experiments to further calibrate their model. [$0.6 \leq \beta < 0.8\bar{3}$, the correlation between $\alpha, \beta$].
- The calibration failed to explain the Public Good Games without Punishment. The discrepancy between the observation and the theory's prediction is well over 30%.
- The calibration was *not* used to explain Gift Exchange Games.

All in all, the only game that F&S have succeeded in explaining is the Market Game with Proposers' Competition (Roth et al. [1991]) in which the responder is restricted to act selfishly. For this game the calibration is superfluous - all distributions explain it.

Thus, the calibration that F&S present in Table III (QJE p. 844) with the additional correlation between the parameters was obtained from the data of the Ultimatum games, the Market Game with Responders' competition and the Public Good Game with Punishment. Given this (coarse) calibration, F&S may now begin to test their theory.

## 4 Competition and Fairness

In this section (section D, QJE p.834) F&S want to demonstrate that the theory of Inequity Aversion agrees with a large body of market experiments which shows that in competitive situations the competitive equilibrium prevails. F&S

want to show that in competitive situations the equilibrium is close to the competitive one even if the population is highly inequity averse.

In their own words: F&S start from the *".. well established experimental fact that in a broad class of market games prices converge to the competitive equilibrium. [Smith 1982, Davis and Holt 1993]."* (QJE p.829), they want to demonstrate that *"convergence to standard competitive predictions can occur even if agents are very strongly concerned about fairness"* (INV p.212).

F&S show it by analyzing two examples of market games and they derive a general principle from these examples. The two games were discussed earlier, one by Roth, Prasnikar, Okuno Fujiwara and Zamir [1991] and the other by Güth, Marchand, and Rulliere [1997]. F&S prove propositions (proposition 2 , 3 QJE pages 830, 832) which describe conditions under which these games have a competitive equilibrium. The two propositions *"suggest that there is a more general principle at work"* (QJE p.834) and this general principle is: *"that no single player can enforce an equitable outcome"* (QJE p.834).

How general can this principle be if derived from two particular examples? How were these examples chosen? In a footnote we are told that the two experiments in this section were chosen for their simplicity and tractability (footnote 9. QJE p. 829). Do these examples possess any other properties except for their simplicity?

We argue that the first game is logically unsuitable to demonstrate their point and that the second game does not show it.

Consider how F&S prove their points.

The first example by Roth et al. is the market game with proposers competition described in section 3.3 above. In this game the responder is restricted to consider only the *highest* offer. This prevents him from practicing any inequity aversion he may have. Consider a responder who received two offers, one which gives him the whole cake and the other half of it. The responder is restricted to choose between accepting the offer 1 and rejecting it and having 0, he is not allowed to consider the offer 0.5. Thus he is prevented from choosing the equitable point, which he may actually prefer. Recall that F&S' aim is to show that the competitive result prevails however *high* the inequity aversion in the population may be. This game does not allow inequity aversion to be arbitrarily high, one specific player is not allowed to have or practice any inequity aversion. The model they chose is simply not suitable for what they wish to demonstrate.

In fact, if we remove this restriction and allow the responder to accept *any* offer (or to reject all offers), and if we endow him with high inequity aversion, $\beta > \frac{n-1}{n}$, then the only equilibrium is the equitable partition in which all the proposers offer $\frac{1}{2}$ and the responder accepts it. Thus the responder can force an equitable outcome irrespective of how selfish the other players are. F&S do not consider this possibility. This shows, contrary to what F&S claim, that fairness considerations do matter even in competitive situations and that their theory (as it is presented in these papers) is not compatible with the experimental observations on competitive markets.

The situation is even more delicate in their second example, the market game with responders' competition, by Güth, Marchand, and Rulliere [1997] discussed in section 3.4 above. In this game a single proposer makes an offer to a number of responders. In Proposition 3 (QJE p.832) F&S show that to guarantee an outcome which is close to the competitive equilibrium the proposer *and* at least one of the responders need to be sufficiently selfish . If the proposer's discomfort at being better off ($\beta$) is not too high, $\beta < \frac{n-1}{n}$, *and* if one of the responders is willing to accept a low share of the surplus then all other responders are forced to accept this share in equilibrium. Thus, how close the equilibrium is to the competitive outcome depends on the most selfish responder *and* the proposer's selfishness.

Like in the previous market game by Roth et al., if the single proposer is sufficiently inequity averse $\left(\beta > \frac{n-1}{n}\right)$ the only equilibrium is one in which he proposes $\frac{1}{2}$ and all accept it. Again, contrary to F&S' general principle, a single player can enforce an equitable outcome.

F&S observe that, as the number of responders $(n-1)$ increases, the proposer is less likely to be sufficiently inequity averse (i.e. with $\beta > \frac{n-1}{n}$). Thus as $n$ increases and the inequality is reversed: $\beta < \frac{n-1}{n}$, the responder prefers the allocation $(1, 0, 0, ..., 0)$ to $\left(\frac{1}{2}, \frac{1}{2}, 0, ..., 0\right)$. This property is a direct consequence of the normalization of the utility function (dividing by $n-1$). The normalization assumption was introduced with no attempt to justify it (QJE p.824), and there seems to be no compelling reason why this consequence should be true. Is it really reasonable that an individual is less likely to contribute to a charity, helping the victims of a natural disaster, just because the number of victims increased? Indeed, tsunamis do not often occur in laboratories but this does not make this consequence more palatable, nor are the populations in the laboratory particularly large to guarantee that any $\beta$ will be smaller than $\frac{n-1}{n}$.

F&S can now deduce their general principle:

*"It is, thus, the impossibility of preventing inequitable outcomes by individual players that renders inequity aversion unimportant in equilibrium."* (QJE p. 834).

It seems that F&S' general principle does not even apply to the two examples from which it was supposed to have been deduced. In the first game the responder who could bring about the equitable outcome was prevented from doing so, and in the second game the proposer, who can force an equitable outcome, is simply being ignored.

The importance of this 'general principle' for F&S is in its use as an introduction to the topic of Endogenous Incomplete Contracts (QJE p.834). F&S modify the market game with proposer competition: they allow the responder to accept *any* offer AND at the same time allow the chosen proposer to withdraw his offer and thus destroy the surplus. This double modification may lead to an equitable outcome. F&S claim that this *"is caused by the fact that in* *the modified game a single proposer can enforce an equitable outcome"* (QJE p.835). They ignore the natural intermediate modification in which the responder is allowed to consider any offer (but the proposer is not allowed to withdraw his offer), which may also lead to an equitable outcome. As a foundation of Endogenous Incomplete Contracts this argument seems rather unsound.

What F&S did in this section is to differentiate between individuals according to their role in the game. The single proposer or the single responder, because of their role in the market, are not allowed to be inequity averse, and are assumed to behave selfishly. F&S do not discuss these point openly, but it is hard to believe that they meant to introduce such differentiation, because it seems to be against the spirit of their theory. If individuals change their behavior (and their preferences) according to the game they play and their role in it, then their behavior is not likely to be explained by a theory which assumes that players have one utility function across games and roles. A theory of social norms will probably do much better. If F&S meant to allow preferences to be dependent on roles, they should have explicitly said so.

It is interesting to see that F&S relinquish all references to the experiment by Güth, Marchand, and Rulliere in INV, they mention only the experiment by Roth et al. As we have seen, the crucial player in Roth et al., the responder, has had his inequity aversion severed and the outcome is the competitive one irrespective of the preferences of the proposers. This way F&S can truthfully (if misleadingly) state that:

> *"For example, in certain competitive environments (see, e.g., the ultimatum game with Proposer competition in Section 5.1) even a population that consists* ***only*** *of very fair types (high $\alpha$ and $\beta$ ) cannot prevent very uneven outcomes."* (INV p.221).

The careful phrasing (*'in certain competitive environments (e.g. Roth et al...)...'*) makes one believe that the authors have a number of such examples. While, in fact, Roth et al. is their only (unsuitable) example, and in a very similar competitive environment (market game with responders' competition) this is no longer true, a single player can enforce an equitable outcome.

# 5 Epilogue

Given this litany of flaws, it was to be expected that F&S would feel the need to provide some explanation of what they have done, if not in these two papers (QJE, INV) then in later papers. However, F&S keep referring to their calibration and citing their inflated description of it in their recent papers.

For example, in a new paper by E. Fehr, A. Klein & K. Schmidt, *"Con-*

*tracts,Fairness, and Incentives"* CESifo Working Paper[10], No. 1215, June 2004, on page 33 they say:

> *"[The assumption is].....derived from the distribution of types calibrated in Fehr and Schmidt (1999) with experimental data on the ultimatum game. Fehr and Schmidt used this distribution to explain the experimental results in many different classes of games, so we want to use it for this game as well."*

But the description of these selfsame results seems to evolves with time. In a recently published paper by Fehr & Schmidt *"Fairness and Incentives in a Multi-Task Principal-Agent Model", Scand. J. of Economics* $106(3), 453 - 474, 2004$ on p. 470, they write (my emphases):

> *"In Fehr and Schmidt (1999), we first calibrated the distribution of the parameters $\alpha_i$ and $\beta_i$ by using the existing experimental evidence on ultimatum and **dictator games** in the literature. Then, we applied the calibrated model to several other classes of games (public good games, market games, **gift exchange games etc.**) and showed that our calibrated model is consistent with the experimental evidence on these games as well."*

No, they did not use the Dictator Game to calibrate their model, nor did they use the calibration to explain gift exchange games. Since they have not even attempted to explain other games in QJE it is not clear what the 'etc.' stands for.

We would like to end on a more positive note: In a recent paper by U. Fischbacher, C. M. Fong and E. Fehr: *"Fairness, Errors and the Power of Competition"* Institute for Empirical Research in Economics, University of Zürich,Working Paper No. 133, December 2003[11], the authors use the distribution of parameters assumed by F&S in QJE.

On p.25 we are told that:

> *"..we completely tied our hands with regard to the choice of the parameters for the FS model by using the same preference parameters as in Fehr and Schmidt (1999) which are presented in Table 1. This is very important because otherwise we could rationalize too many different outcomes by choosing the appropriate preference parameters."*

[10] The paper was published as a CESifo discussion paper and can also be downloaded from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=549224

[11] It is natural for discussion papers to change over time. The version I looked at was downloaded from http://www.iew.unizh.ch/wp/iewwp133.pdf on 25.2.2005.

The authors (E. Fehr is one of the authors but not K. Schmidt) admit that too much freedom in choosing a model's parameters is not a rigorous scientific method. But what is it they tie their hands to? The only information FFF give us about the distribution in Table 1, which they borrow from QJE, is that it was assumed by F&S, no other reason or justification for using it is offered in the paper's 27 pages. They wish to tie their hands for scientific reasons, but they should also provide a good reason why they chose this element to be bound to.

We will have to wait and see what the future will bring.

# 6 Conclusion

It seems that at least in part of the field of experimental economics, the scientific standards may not be as rigorous as one may desire. It was in the name of the sophisticated and rigorous scientific techniques of experimental economics that F&S asked us to accept their inequity aversion utility function (INV p.209). Can we trust their judgement? Their own practices cast a dark shadow on the theory of inequity aversion. Is the theory so weak as to require such aggressive marketing techniques?

But the sad point is not only that two distinguished economists allow themselves to write in this style, it is also that they managed to get by, that these papers have caused no outcry or protest in the experimentalist community[12]. Could it be that, with the exception of QJE's proofreader, no one has ever carefully read the details of the paper until now?

Experimental Economics has the great advantage that, in a superficial level, the details of the experiment are easy to understand and are readily enjoyable. Almost anyone, whether a professional or a layman, is an amateur psychologist and a keen observer of human nature. Anyone will therefore have an opinion about the implications of an experiment, how it can be bettered and what the outcome should be. Compare the lulling effect of a typical JET or Econometrica paper, on an audience of non economists, with the enlivening effect that a good experiment can have on a dinner party.

In comparison with the appeal of the narrative part, the scientific part of an experimental study is rather dull. The minute analysis of the data, the statistical details, the intricate comparison with the details of other experiments are all necessary and important, but hardly inspiring.

This difference in interest and appeal of the narrative part and that of the analytical parts, is probably larger in experimental economics than in any other economic field.

It is therefore very tempting, when reading an experimental paper, to read only the first anecdotal part. After all, any good economist can tell from this

[12] There were, and still are, debates about the theory's relevance to efficiency and 'intentions'. There seem to be no discussions of the validity of F&S' claims.

part what the implications of the results to economic reasoning are and in which direction the research should proceed. There is an understandable tendency to neglect the duller scientific details of the paper.

I argue that this a dangerous policy which may adversely distort our decisions, decisions about promotions and appointments of colleagues, of accepting a paper for publication and of accepting a 'tested' theory. I plead for more care.

# 7 Appendix

To aid the reader we reproduce here some tables, figures, equations and a proposition which may help the reader follow the details of the arguments in this pamphlet.

TABLE I
PERCENTAGE OF OFFERS BELOW 0.2 AND BETWEEN 0.4 AND 0.5 IN THE ULTIMATUM GAME

| Study (Payment method) | Number of observations | Stake size (country) | Percentage of offers with $s < 0.2$ | Percentage of offers with $0.4 \leq s \leq 0.5$ |
|---|---|---|---|---|
| Cameron [1995] (All Ss Paid) | 35 | Rp 40.000 (Indonesia) | 0 | 66 |
| Cameron [1995] (all Ss paid) | 37 | Rp 200.000 (Indonesia) | 5 | 57 |
| FHSS [1994] (all Ss paid) | 67 | $5 and $10 (USA) | 0 | 82 |
| Güth et al. [1982] (all Ss paid) | 79 | DM 4–10 (Germany) | 8 | 61 |
| Hoffman, McCabe, and Smith [1996] (All Ss paid) | 24 | $10 (USA) | 0 | 83 |
| Hoffman, McCabe, and Smith [1996] (all Ss paid) | 27 | $100 (USA) | 4 | 74 |
| Kahneman, Knetsch, and Thaler [1986] (20% of Ss paid) | 115 | $10 (USA) | ? | 75[a] |
| Roth et al. [1991] (random payment method) | 116[b] | approx. $10 (USA, Slovenia, Israel, Japan) | 3 | 70 |
| Slonim and Roth [1997] (random payment method) | 240[c] | SK 60 (Slovakia) | 0.4[d] | 75 |
| Slonim and Roth [1997] (random payment method) | 250[c] | SK 1500 (Slovakia) | 8[d] | 69 |
| Aggregate result of all studies[e] | 875 | | 3.8 | 71 |

a. percentage of equal splits, b. only observations of the final period, c. observations of all ten periods, d. percentage of offers below 0.25, e. without Kahneman, Knetsch, and Thaler [1986].

**Table I, QJE p.827**

TABLE II
PERCENTAGE OF SUBJECTS WHO FREE RIDE COMPLETELY IN THE FINAL PERIOD OF A REPEATED PUBLIC GOOD GAME

| Study | Country | Group size (n) | Marginal pecuniary return (a) | Total number of subjects | Percentage of free riders ($g_i = 0$) |
|---|---|---|---|---|---|
| Isaac and Walker [1988] | USA | 4and10 | 0.3 | 42 | 83 |
| Isaac and Walker [1988] | USA | 4and10 | 0.75 | 42 | 57 |
| Andreoni [1988] | USA | 5 | 0.5 | 70 | 54 |
| Andreoni [1995a] | USA | 5 | 0.5 | 80 | 55 |
| Andreoni [1995b] | USA | 5 | 0.5 | 80 | 66 |
| Croson [1995] | USA | 4 | 0.5 | 48 | 71 |
| Croson [1996] | USA | 4 | 0.5 | 96 | 65 |
| Keser and van Winden [1996] | Holland | 4 | 0.5 | 160 | 84 |
| Ockenfels and Weimann [1996] | Germany | 5 | 0.33 | 200 | 89 |
| Burlando and Hey [1997] | UK,Italy | 6 | 0.33 | 120 | 66 |
| Falkinger, Fehr, Gächter, and Winter-Ebmer [forthcoming] | Switzerland | 8 | 0.2 | 72 | 75 |
| Falkinger, Fehr, Gächter, and Winter-Ebmer [forthcoming] | Switzerland | 16 | 0.1 | 32 | 84 |
| Total number of subjects in all experiments and percentage of complete free riding | | | | 1042 | 73 |

**Table II, QJE p.838**

Free riding in public good games with no punishment

$$
(5) \qquad s^* \begin{cases} = 0.5 & \text{if } \beta_1 > 0.5 \\ \in [s'(\overline{\alpha}), 0.5] & \text{if } \beta_1 = 0.5 \\ \in (s'(\underline{\alpha}), s'(\overline{\alpha})] & \text{if } \beta_1 < 0.5. \end{cases}
$$

**Equation 5, QJE p.827**

The optimal offers ($s*$) in an ultimatum game as a function of the inequity aversion parameter ($\beta$). Note that for those proposers who offer 0.5 all that the theory specifies is $\beta \geq 0.5$.

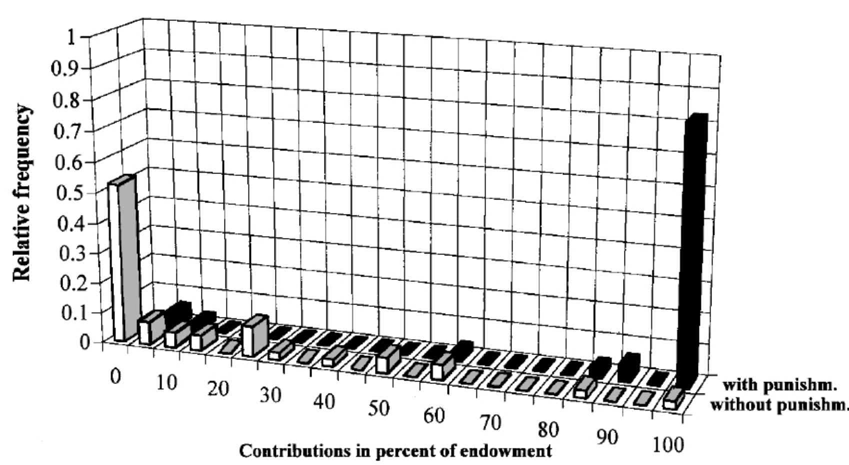


FIGURE II

Distribution of Contributions in the Final Period of the Public Good Game with Punishment (*Source:* Fehr and Gächter [1996])

**Figure II, QJE p.839**

TABLE III

Assumptions about the Distribution of Preferences

| Distribution of $\alpha$'s and associated acceptance thresholds of buyers | | | Distribution of $\beta$'s and associated optimal offers of sellers | | |
|---|---|---|---|---|---|
| $\alpha = 0$ | 30 percent | $s' = 0$ | $\beta = 0$ | 30 percent | $s^* = 1/3$ |
| $\alpha = 0.5$ | 30 percent | $s'(0.5) = 1/4$ | $\beta = 0.25$ | 30 percent | $s^* = 4/9$ |
| $\alpha = 1$ | 30 percent | $s'(1) = 1/3$ | $\beta = 0.6$ | 40 percent | $s^* = 1/2$ |
| $\alpha = 4$ | 10 percent | $s'(4) = 4/9$ | | | |

**Table III (QJE p. 844)**
The calibration

# References


Fehr, E., and S. Gächter (2000): "Cooperation and Punishment in Public Goods Experiments," *American Economic Review*, pp. 980 – 994.

Fehr, E., A. Klein, and K. M. Schmidt (2004): "Contracts, Fairness and Incentives," CESIFO Discussion Paper, No. 1215.

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

——— (2003): "Theories of Fairness and Reciprocity: Evidence and Economic Applications," in *Advances in Economic Theory, Eigth World Congress of the Econometric Society*, ed. by S. T. M. Dewatripont, L.P. Hansen, vol. Vol. 1, pp. 208–257. Cambridge University Press, Cambridge, Vol. 1.

——— (2004): "Fairness and Incentives in a Multi-Task Principal-Agent Model," *Scand. J. of Economics*, 106(3), 453 – 474.

Fischbacher, U., C. Fong, and E. Fehr (2003): "Fairness, Errors and the Power of Competition," Institute for Empirical Research in Economics, University of Zurich, Working Paper No. 133.

Güth, W., N. Marchand, and J.-L. Rulliere (1997): "Ultimatum Bargaining Behavior - A Survey and Comparison of Experimental Results," Discussion Paper, Humboldt University, Berlin.

Roth, A. E. (1995): "Bargaining Experiments," in *Handbook of Experimental Economics*, ed. by J. Kagel, and A. Roth. Princeton University Press.

Roth, A. E., V. Prasnikar, M. Okuno-Fujiwara, and S. Zamir (1991): "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study," *American Economic Review*, LXXXI, 1068 – 1095.