An ESRC Research Group

Inequality Aversion and Stochastic Decision-making: Experimental Evidence from Zimbabwean Villages after Land Reform

GPRG-WPS-061

Stefan Kohler

Global Poverty Research Group

Website: http://www.gprg.org/

The support of the Economic and Social Research Council (ESRC) is gratefully acknowledged. The work was part of the programme of the ESRC Global Poverty Research Group.

Inequality Aversion and Stochastic Decision-making: Experimental Evidence from Zimbabwean Villages after Land Reform

Stefan Kohler*

November 27, 2006

Abstract

Inequality considerations are a motive for making positive offers in the Ultimatum Game and rejecting small ones, but decision error could have the same effect. I find evidence for both of these considerations and a different relative importance amongst Zimbabwean villagers, of whom some resettled after a government organized land reform during the 1980s. Resettled villagers have higher inequality aversion and lower decision error than those who live in traditional villages but, after accounting for different levels of inequality aversion, the difference in decision error between both groups of villagers is no longer significant. There are no gender differences in preferences. The model estimated was first used by De Bruyn and Bolton (2004) on a large set of bargaining data but the best fit of 64 percent overall coincidence of observed and predicted behavior is achieved for a different 'symmetric' specification of inequality aversion in the model. As the use of field data is a recent development in experimental economics, I reestimate the model applied to the Zimbabwean data on the laboratory Ultimatum Game data of Roth et al. (1991) and further field data from Henrich et al. (2005). Estimates are compared comprehensively.

JEL classification: C72, C93

Keywords: Quantal Response Equilibrium, Ultimatum Bargaining Game, Inequality Aversion, Gender Difference, Land Reform, Field Experiment

*European University Institute, Department of Economics, Villa San Paolo, Via Piazzuola 43, 50133 Florence, Italy. E-mail adress: stefan.kohler@eui.eu

1 Introduction

In laboratory experiments of the Ultimatum Game the modal offer is typically 50 percent, and the mean offer lies between 40-50 percent of total. Responders usually accept average offers but reject offers lower than 20 percent of total. (Camerer and Thaler 1995, Roth 1995) These deviations are robust with respect to stake size, context, and within similar countries, but, as for instance, Henrich (2000) shows Machiguena (Peru) behave very differently from subjects drawn from industrialized populations. In this paper game behavior from a field experiment in rural Zimbabwe is studied. Ultimatum Games were played in 20 villages, some of them resettled after a land reform. Overall, in the play amongst villagers the familiar principal pattern emerges: (R1) In every village observed bargaining outcomes are significantly different from the Subgame Perfect Equilibrium prediction. (R2) Some rejections of positive offers result in Pareto inefficient outcomes. I further find (R3) no gender difference in bargaining behavior and (R4) on average higher offers in resettled villages, while (R5) the rejection rates are the same across resettled and traditional villages types.

Some studies have sought to explain the regularities (R1) and (R2) for the Ultimatum Game by assuming subjects' preferences depend not only on their own monetary payoff but also on others' payoff in various ways, so called *social preferences*, as proposed in general models by Rabin (1993), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Rabin and Charness (2002) or Kohler (2005). Others have sought to explain the results without social preferences, assuming expected monetary payoff maximization, but studying adaptive learning dynamics (Prasnikar and Roth, 1992), reinforcement learning model (Roth and Erev, 1995), evolutionary dynamics (Gale et al., 1995), limited cognition (Johnson et al., 2002), or decision error (Yi, 2005) in the notion of McKelvey and Palfrey's (1995, 1998) Quantal Response Equilibrium.

This paper follows a combined approach. I estimate a social preference of inequality aversion, as suggested by Bolton and Ockenfels (2000), in a Quantal Response Equilibrium as a model of error prone behavior. Decision error is caused by a random shock to experienced utility converting the deterministic to a stochastic choice model. The stochastic model is estimated with standard likelihood methods after making a distributional assumption on the shock. Estimating the preference parameters, rather than merely calibrating them, which is usually done in social preference studies,[1] allows me to statistically assess how well the Zimbabwean Ultimatum Game behavior can be described by the social preference

[1] See, for instance, the aforementioned Rabin (1993), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Rabin and Charness (2002) and Kohler (2005).

and decision error. The model employed for the analyses is similar in spirit to Goeree and Holt (2000) and borrowed from De Bruyn and Bolton (2004), who assess the plausibility of the decision error and inequality aversion approach in bargaining situations by testing its out-of-sample forecast accuracy between different experimental datasets from industrialized populations.[2]

I complement the previous work by using field data from Ultimatum Game experiments, which were carried out across villages in Zimbabwe. Unlike the laboratory data used by De Bruyn and Bolton, the Zimbabwean field data is better explained by a model, which is symmetric rather than asymmetric in inequality aversion. Both parameters, inequality aversion and decision error, are significant at the 1 percent level and the model predictions coincide on average with 64 percent of the observed behavior. The standard game theoretic solution with pure self-interest predicts less than 2 percent correctly. Hence, reestimation of De Bruyn and Bolton (2004)'s model confirms the relevance of decision error and social preferences as well as the predictive success of a flexible social preference model in a distinct economic environment. Splitting the Zimbabwean sample, I do not observe any gender difference in the Ultimatum Game behavior and estimated parameters, but villagers, who have resettled within a land reform, exhibit higher inequality aversion. A regression analysis of the game data (see also Barr, 2004a) reveals similar differences, but does neither link them to equilibrium behavior nor allows for a quantitative analysis of the different play with respect to context independent behavioral motives/incentives. To my knowledge, this paper is the first one to analyze gender difference or the incentive structure after a policy intervention with a structural social-preference model.

Furthermore, as generally the use of field data is a novum in experimental economics and no decision error model has previously been estimated with field data, I investigate if there are systematic differences between field and laboratory play, which can be attributed to inequality aversion and/or decision error. Henrich et al. (2005) in a seminal paper show, by bringing experimental games as the Ultimatum Game, Public Good Game, or Dictator Game into non-industrialized societies, that differences in behavior across societies were substantial and influenced by the local norms and traditional social structures. Reanalyzing part of their data and of Roth et al.'s (1991) four-country bargaining study, I find that decision error increases in the field despite the fact of a higher amount of money at stake.[3]

The paper is structured as follows. Section 2 reviews related literature. Section 3 de-

[2]The model predicts well across distinct classes of bargaining games, ranging from Ultimatum to alternating-offer bargaining games.

[3]Roth et al. (1991)'s data is the Ultimatum Game bargaining data analyzed by De Bruyn and Bolton (2004). As this paper uses a different model I do not refer to their estimates.

scribes the dataset and gives background information on the Zimbabwean context. In section 4, I describe the Quantal Response Equilibrium concept and the model. Section 5 reports on the estimation results and model performance. I also compare alternative model specifications. The findings from the Zimbabwean field experiment are then contrasted with other data in section 6. Section 7 concludes.

2 Related Literature

The idea to structurally estimate not only decision error but also social preference parameters in a Quantal Response Equilibrium (henceforth QRE) model was pioneered by Palfrey and Prisbrey (1997), who estimate altruism in a probit model on contributions in a binary choice Public Good Game.[4] Estimation of social preferences in a QRE was then consequently pursued by Anderson et al. (1998), Offerman et al. (1998), Goeree and Holt (2000), and lately by aforementioned De Bruyn and Bolton (2004). All but the last two groups of authors investigate decision error and altruism. Goeree and Holt (2000), in contrast to the latter group, estimate the inequality aversion parameters of the Fehr and Schmidt (1999) social preference specification. Costa-Gomes and Zauner (2001) estimate altruism and decision error in a Bayesian Nash Equilibrium of the Ultimatum Game (henceforth UG).

All these authors, but Costa-Gomes and Zauner (2001), have applied the QRE as defined by McKelvey and Palfrey (1995, 1998) for normal- or agent normal-form games, respectively. Thereby, both notions of the QRE describe well patterns of deviations from equilibrium observed in a variety of normal- and extensive form experiments. Since, for instance, the UG has a normal- and agent normal-form representation in principle both notions of the QRE can be applied. Yi (2005) studies both, the normal form QRE (henceforth NQRE) and agent QRE (henceforth AQRE), on the UG data of Slonim and Roth (1998). He finds that coordination patterns in the repeated UG play of this experiment is better explained in a NQRE, as "the existence of multiple NQRE make it possible that equilibrium selection is influenced by the history of play." (p. 340) In contrast, the sequential rationality imposed in the QRE of the agent-normal form implies that the AQRE converges to the unique subgame-perfect equilibrium with the smallest feasible strictly positive offer as bargainers make fewer decision errors.

For the primary purpose of empirically comparing the two notions of QRE Yi (2004, p. 334) decided to disregard aspects of social preferences and emphasizes that "the main issue is

[4] Palfrey and Prisbrey (1997) find a fair amount of noise in the decision process and significant warm-glow effects (albeit low in magnitude). They do not find evidence of pure altruism.

not whether social utility is necessary to explain the data." Instead, his analysis "focuses on the relevance of backward induction that underlies AQRE along the same lines of the recent experimental studies as Binmore et al. (2002) and Johnson et al. (2002) who showed that the subjects' behavior in various bargaining games consistently violates subgame consistency even while controlling for social utility." As subgame perfection is not debated in this paper and the Ultimatum Games considered are one-shot, requiring sequential rationality for the equilibrium play of the UG seems a preferable choice to the author. I follow the standard approach for extensive form games by applying the AQRE to study the Zimbabwean UG behavior. Any further mention of the QRE therefore refers to its AQRE notion for agent normal-form games if not made explicit otherwise.

3 Data and Background

In 1999/2000 the Zimbabwean government was starting to prepare for a second wave of organized land reform. A first resettlement exercise was undertaken in the early 1980s after the replacement of the Rhodesian government, when land previously owned by white farmers was reallocated to indigenous households. "In response the Southern African Regional Institute for Policy Studies (SARIPS), with the assistance of several international donors, was setting up the Zimbabwe Land Reform Research Network (ZLRRN). The aim of this research network was to look at the lessons that could be learnt from the earlier resettlement exercise, as to derive insights about the process by which social capital is created and about its importance for the process of development more generally." (Barr, 2002, pp. 1-2)

In that context socioeconomic data was gathered from 1999 to 2001 within the yearly Zimbabwe Rural Household Dynamics Project (ZRHDP) survey in 22 resettled villages and six non-resettled villages. At the same time the UG, Trust Game, Public Good Game and Risk Pooling Game experiments were run by Barr (2004a, 2004b, 2003a, 2003b, 2002, 2001) in randomly selected households in some of those villages.[5] The possibility to combine these data from different sources yields an exceptionally rich dataset. Data collected in surveys typically is neither linked to, nor does contain experimental components. In turn, experimental economists usually do not collect comprehensive socioeconomic information from their subject pools. Fehr et al. (2003) argue about the potential of such datasets and suggest methods how economic experiments can be included in a paper based survey.

Another particular aspect of the experimental data from Zimbabwe is that economic

[5] In the experiments the sampling proportion varied between 0.31 and 0.59 or was 1.

experiments were carried out in the field; i.e. with a sample of a village population.[6] Typically, in 'laboratory' experiments there is less heterogeneity amongst the participants, who are often all students. According to Barr (2002), to ensure comparability between the field data and existing UG data, great care was applied to resemble laboratory conditions. Most importantly, play was anonymous and interaction finite. In the field, play was for realistic stakes between a half and two day's casual labour wage. Since the subject pool was small and experimentees always knew at the start of a game that they were playing with someone from their village, the players' "expectations about others' behavior [...] was always based on their past experiences." (Barr 2002, p. 13) Thus, in the field experiments a prior belief about the other players' action incorporates knowledge about local social norms and behavior.

In the paper at hand the UG data from the above dataset is used. The UG is a very simple strategic situation. Two payers are allotted a sum of money (termed the pie or stakes) and then bargain about its division. The first player, called proposer, makes an offer, which the second player, the responder, can accept or reject. If the responder accepts then the pie is divided according to the proposed split. If the responder rejects, both receive nothing. Positive theory, especially Subgame Perfect Equilibrium and money maximization, predict that proposers should offer the smallest non-zero amount and responders should always accept, because they face a choice between zero and something. (In the subgame perfect case 0 and 1 unit.) In contrast, typical experimental UG behavior and the behavior observed in all Zimbabwean villages substantially deviates from this prediction of game theory under standard preferences.

Table 1 below summarizes the mean and modal offers, the mean acceptance probability, and the rejection rate of low offers (below 20 percent) from the Zimbabwean UG, which was played in 20 villages. It also splits up these statistics by the resettlement status of the villages and by gender. Participant households for the UG experiment were invited randomly within villages. According to Barr (2004a, p. 313) "each household was asked to send one adult (above the age of 14). The chairman or headman of each village was charged with the duty of ensuring [...] that between forty and sixty percent of the volunteers were women."[7] The villages represent a stratified sample, in which attention was paid to cover to an equal extend villages from different areas with a fixed proportion of resettled and traditional villages. Tables A1 (p. 37) and A2 (p. 38) in the appendix A provide first

[6] For a discussion on the implementation procedure of field experiments in developing countries see Barr (2003) or Henrich (2000).

[7] Thus, there might be an issue of self-selection bias that cannot be not tackled using the available data.

moments of the data on the village level and some comparison of resettled and traditional villages.

Table 1: *UG Descriptive Statistics*

	all	resettled	traditional	female	male
N	242	180	62	111	131
mean offer in percent	0.44	0.45	0.41^{Δ}	0.44	0.44
	(0.12)	(0.10)	(0.14)	(0.11)	(0.12)
modal offer in percent	0.50	0.50	0.50	0.50	0.50
	(155/242)	(121/180)	(34/62)	(70/111)	(85/131)
acceptance rate	0.92	0.93	0.90	0.92	0.92
	(223/242)	(167/180)	(56/62)	(102/111)	(121/131)
low offer rejection rate	0.43	0.50	0.34	0.50	0.38
	(12/28)	(8/16)	(4/12)	(8/16)	(6/16)

Note: Standard deviations or actual numbers are given parentheses. Max. offer Zim\$ 50. Low offer is 20 percent or less. $^{\Delta}$Indicates significant difference at 1 percent/5 percent level using a Student-t/Epps-Singleton test. No further significant difference at the 1, 5 and 10 percent levels.

The stakes in the administered UG were Zim\$ 50 and the smallest unit to offer was Zim\$ 5.[8] All mean offers are between 40 and 50 percent and the modes are all 0.50. A modal offer is unanimously accepted with certainty, but offers lower than 20 percent of the total are frequently rejected. These stylized findings from the field experiment are in line with the typical experimental evidence from laboratory experiments.[9] In addition, 14 percent (1 out of 7 responders) rejected a high offer of Zim\$ 30. Student-t and Epps-Singleton[10] tests to test for different proposer and responder behavior across resettlement status and gender, support no difference in behavior across gender and similar responder behavior between resettled and traditional villages, but reject that mean offers in resettled and traditional villages are the same. Appendix A tables A3 (p. 39) and A4 (p. 40) further contain

[8] This led to average earnings between half a day's and a day's casual wage.

[9] Summarized, for instance, by Camerer and Thaler (1995) or Roth (1995).

[10] The true distribution underlying the observed UG behavior is unknown. Student-t tests are appropriate if the distribution is continuous. Epps-Singleton (1986) test are for discrete distributions. The results of both tests are the same at the 5 percent level.

regression results which, overall, support this view.[11]

4 Stochastic Choice, Social Preferences and Individual Decision Making

4.1 The Quantal Response Equilibrium

4.1.1 Intuition

McKelvey and Palfrey (1995, 1998) investigate how players choose best strategies in an environment, in which they make decision errors. More precisely, players choose among strategies in a game based on randomly disturbed (or noisy) expected utility. A random shock, which adds to the expected utility, causes players to make errors in the sense of not always choosing their best strategy given their beliefs.[12] Each player knows that the other players do so as well and anticipates the decision error of others. For a given error structure a QRE[13] is defined as a fixed point of this process. McKelvey and Palfrey establish the existence of this fixed point under the behavioral assumptions that players ($Q1$) maximize utility, and ($Q2$) estimate expected payoffs in an unbiased way. With these assumptions individual choices remain rational in the QRE but are based on latent variables, which are the player's vector of estimated payoffs.

McKelvey and Palfrey (1995)'s original notion of QRE is a normal-form concept.[14] The NQRE gives identical predictions for extensive- and normal-form representations of a given game. However, experimental subjects behave "systematically different[ly] in normal- and extensive-form representations of many games, including ultimatum bargaining games (Schotter et al., 1994; Cooper and van Huyck, 2002). In response to these difficulties, McKelvey and Palfrey (1998) extended the NQRE to extensive-form games, proposing a notion

[11] However, simple regression yields biased coefficients and invalid hypothesis testing if responder behavior depends on anticipated proposer action and vice versa. Simultaneous regression techniques are then more appropriate.

[12] According to Goeree et al. (2004, p. 3), "stochastic choice via a mixed strategy as a solution to an optimization problem [QRE] contrasts sharply with the introduction of decision errors in the analysis of refinements, as in Selten (1965) and Myerson (1978)." Latter are "based, ironically, on the notion that players do not always maximize, and hence reasonable equilibria are the ones that are robust to 'trembles' by the opponent. Also in contrast with mixed strategies these error-based refinements are defined in terms of asymptotic properties when the errors become negligible."

[13] "The name is borrowed from the statistical literature on quantal choice/response models in which individual choices are rational, but are based on latent variables." (McKelvey and Palfrey, 1995, p. 7)

[14] See appendix B for a formal definition of the NQRE.

they called AQRE." (Yi 2005, p. 325)

The AQRE is defined like a NQRE, but for the agent normal-form of an extensive-form game, in which different information sets of a given player are assumed to be played independently by different agents, but all of a given players's agents share the same payoff function. "Because each agents's noise is assumed to be independent, for any game with a non-trivial extensive form, an AQRE differs from a NQRE, where the noise terms for the agents of a given player are in effect assumed to be perfectly correlated." (Yi 2005, p. 325)

In consequence of the noisy utility, players do not always choose their best response. The best response functions and equilibrium behavior in any QRE become stochastic rather than deterministic (at least for an outside observer) and all strategies of a player can be observed with some probability in equilibrium. Because the QRE makes a stochastic equilibrium prediction, in contrast to the deterministic Nash equilibrium prediction, experimental data can be used to obtain maximum likelihood estimates of the decision error and utility parameters.[15] When applying a QRE model, the random utility shocks and the thereby implied strategy choices follow a specific distribution. Assuming the best responses functions follow a logit distribution the specific QRE is often called a logit equilibrium and the decision error can be represented by just one distributional parameter. The logit response function corresponds to optimal choice behavior in a QRE if ($Q3$) for each player (or agent) i and strategy j the error distribution f_i has an extreme value distribution with cumulative density function $F_i\left(\varepsilon_{ij}\right) = e^{-e^{-\lambda\varepsilon_{ij}-\gamma}}$ and the ε_{ij}'s are independent. In the logit equilibrium the choice probability for each player i and strategy j is given by:

$$\sigma_{ij}\left(E\left[u_{ij}\right]\right) = \frac{e^{\lambda E[u_{ij}]}}{\Sigma_{k=1}^{J_i} e^{\lambda E[u_{ik}]}}, \tag{1}$$

where $E\left[u_{ij}\right]$ is the expected utility of player i's strategy j.

The logit response function implies that strategies with higher expected payoffs are chosen with higher probability. McKelvey and Palfrey (1995, 1998) show that the logit equilibrium approaches a subset of Nash equilibria as noise disappears.[16] This alternative approach "does not abandon the notion of equilibrium, but instead replaces the perfectly rational expectations equilibrium embodied in Nash equilibrium with an imperfect, or noisy, rational expectations equilibrium." [...] The QRE model is therefore a "natural extension of well-developed and commonly used statistical models of choice or quantal response that have a long tradition in statistical applications to biology, pharmacology, and the social sciences." (McKelvey and Palfrey 1995, p.7)

[15] Parameters are identified if they have independent variation in the data.

[16] This finding is restricted to games with discrete, finite strategy spaces in McKelvey and Palfrey (1995) and extended to contineous, finite strategy spaces by Yi (2005).

4.2 Decision Making in the Ultimatum Game

In the following I apply the logit equilibrium (1) to the agent normal-form of the UG. In particular, I study UG behavior under the preference assumptions ($P1$) that players have a social preference as suggested by Bolton and Ockenfels (2000), and ($P2$) that a representative agent is sufficient to understand the essential UG behavior. Explicitly ($P2$) implies that there is no systematic difference in the social preference and the distribution of the decision error amongst players within specific groups. Together assumptions ($P1$), ($P2$) and ($Q1$)-($Q3$) exhaustively describe the model. I perform a limited test of the assumed functional form ($P1$) on the Zimbabwean UG data later.

To solve for the logit equilibrium of the UG, I define the game in its agent normal-form formally. Let i denote a player, $i \in \{P, R\}$, where P identifies the proposer and R the responder. The proposer chooses strategy $s \in S_P$, where s denotes how much the proposer gives to the responder and $1 - s$ is what he would like to keep for himself. Each of the responders agents $R(s)$ is identified by the offer. In experiments $S_P = \left\{0, ..., \frac{c}{m}, ..., c\right\}$ is discrete. Constant c represents the pie size and $m > 1$ a finite integer. J_P is the number of pure strategies for the proposer. The responder's behavioral (or his agent's) strategy $r(s) \in S_R$ is a function that maps each possible offer into $S_R = \{accept, reject\}$. In UG experiments the responder can observe the proposers offer before he makes his decision. The AQRE logit equilibrium imposes sequential rationality by requiring that the responder evaluates the observed information when making his decision. That is, the responder's choice is conditional on the observed offer and therefore, unlike the proposer's choice, under certainty. For the agent normal-form of the UG, the logit response functions of the proposer and responder are derived under general preference assumptions below, after a discussion of the social preference employed in the estimation. The section then ends stating the likelihood function and discussing identification of the decision error and social utility parameter of the model.

4.2.1 The Social Preference Function

In an inequality aversion model bargainers have a preference for absolute (monetary) and relative (fairness) payoff. The relative payoff is a loss of utility whenever bargaining outcomes are unequal. The attractiveness of inequality aversion models is based on their ability to rationalize a number of well known anomalies, amongst which is the bargaining behavior in the UG, with just two motives, maximization of own payoff and inequality concern.

Explicitly, each player is assumed to have symmetric social preferences of inequality aversion (henceforth ERC) as suggested by Bolton and Ockenfels (2000):

$$u_i(x_i, x_j) = \begin{cases} c\left(\omega_i - \frac{b}{2}\left(\omega_i - \frac{1}{2}\right)^2\right) & \text{if} \quad x_i, x_j \neq 0 \\ 0 & \text{if} \quad x_i, x_j = 0 \end{cases}, \tag{2}$$

where $\omega_i = \frac{x_i}{c}$ is the proportion of the pie player i gets. Parameter b measures the relative importance of relative gains. It can be interpreted as concern for fairness or inequality aversion. Absolute and relative payoffs are additively separable. The relative payoff is an asymmetric loss function, with minimum loss zero when the player receives half of the pie. The marginal loss from an unequal payoff distribution is greater the further the player's share is below or above half.

This utility function has one parameter b to be fitted. It will be interpreted as a population average. As the random utility shocks ε in the model cause different experienced utility at identical behavior, one interpretation of decision error in the QRE framework is that it represents unmodelled/unobserved heterogeneity between individuals, which is not covered by the preference parameter b. For notational convenience, from now on I denote $u(1-s) := u_P(1-s, s)$ the proposer utility and $u(s) := u_R(s, 1-s)$ the corresponding responder utility if an offer is accepted. If the responder rejects an offer in the UG then both players receive zero. This outcome $u_i(0,0) = 0$ is denoted $u(\varnothing)$ for both players $i \in \{P, R\}$.

4.2.2 The Responders Decision

For a given agent of the responder, the logit response function maps his (expected) utility from each possible pure strategy $r(s) \in S_R$ into a mixed strategy σ_R. In this mixed strategy, let $\sigma_R(s)$ be the probability that the responder R accepts an offer of proportion s of the pie:

$$\sigma_R(s) = \frac{e^{\lambda u(s)}}{e^{\lambda u(\varnothing)} + e^{\lambda u(s)}} = \frac{e^{\lambda u(s)}}{1 + e^{\lambda u(s)}} \tag{3}$$

Responders' probability to accept an offer of s

This probability $\sigma_R(s)$ follows directly from the general logit response function (1) and the responders strategy set $S_R = \{accept, reject\}$. As the responder is the second mover in the UG, there is no uncertainty about the offer and his (expected) utility from accepting an offer s is $u(s)$. The utility of rejection $u(\varnothing)$ is zero. Therefore, the term $e^{\lambda u(\varnothing)}$ equals one and the denominator can be simplified. By definition $\sigma_R(s) \in [0,1]$ and $\sum_{r(s) \in S_R} \sigma_R(s) = 1$.

For the responder, $\lim_{\lambda\to\infty}\sigma_R(s) = 1$ if $u(s) > u(\varnothing)$ and $\lim_{\lambda\to\infty}\sigma_R(s) = 0$ if $u(s) < u(\varnothing)$. Further, $\lim_{\lambda\to 0}\sigma_R(s) = 0.5$. That is, the higher the coefficient of certitude λ the more likely the responder chooses the strategy that yields the highest utility. At the other extreme if the responder is completely irrational then he is indifferent between accepting and rejecting, independent of the actual value of s. That is, his decisions are arbitrary.

4.2.3 The Proposers Decision

Also the proposer's decision to offer s, and to keep $1-s$ for himself, follows a logistic distribution. Let σ_P be a mixed strategy for the proposer. Then the probability, $\sigma_P(s)$, in this mixed strategy that the proposer P makes any particular offer s of the pie from his strategy set $S_P = \left\{0, ..., \frac{c}{m}, ..., c\right\}$ is given by:

$$\sigma_P(s) = \frac{e^{\lambda E[u(1-s)]}}{\sum\limits_{s\in S_P} e^{\lambda E[u(1-s_j)]}} = \frac{e^{\lambda\sigma_R(s)u(1-s)}}{\sum\limits_{s\in S_P} e^{\lambda\sigma_R(s)u(1-s_j)}}, \tag{4}$$

Proposers' decision rule; probability to make an offer of s

where $E[u(1-s)]$ is the expected utility of offering s to the responder. The probability $\sigma_P(s)$ follows directly from the general logit response function (1) and the proposers strategy set S_P. Consistent with assumption $(Q2)$ I suppose that proposers will anticipate the probability that responders will accept any particular offer correctly. Hence, $E[u(1-s)] = \sigma_R(s)\,u(1-s)$. By definition $\sigma_P(s) \in [0,1]$ and $\sum_{s\in S_P}\sigma_P(s) = 1$. Since $\partial\sigma_P(s)/\partial E[u(1-s)] > 0$, offers with higher expected utility are likely to be chosen more often. As the coefficient of certitude increases towards infinity, the proposer systematically makes the offer s that induces the highest expected utility; i.e. $\lim_{\lambda\to\infty}\sigma_P(s) = \arg\max_{s\in S_P}\{E[u(1-s)]\}$. As certitude disappears all offers are made with equal probability; i.e. $\lim_{\lambda\to 0}\sigma_P(s) = 1/J_P$ for all $s \in S_P$.

Due to assumption $(Q3)$ the logit response functions $\sigma_P(s)$ and $\sigma_R(s)$ define the logit equilibrium for the UG. It is noteworthy that, when applying the logit equilibrium to the UG with discrete strategy spaces, the multinomial logit specification implies independence of irrelevant alternatives (see, for instance, Green 2003, p. 724). That is expected utilities of each strategy should be independent of one another. As expected utilities are point estimates of an underlying continuous utility function this assumption is violated. Altering the number choices, "without modifying the size of the pie, would add artificial noise." (De Bruyn and

Bolton, 2004, p. 16) This 'artificial noise' makes it impossible to compare estimates from studies, in which the proposer has a different number of alternatives, or to estimate one model (a unique set of parameters) across such data. De Bruyn and Bolton (2004), who compare decision error across different studies and game specifications, compensate for artificial noise by multiplying the decision error by the log of the number of choice options. Within most of the estimation in this paper, I do not encounter this problem as I restrict the estimation of parameters to Ultimatum Games, which all give the same number of choices to both players ($J_P = 11, J_R = 2$). Only when comparing the estimates from the field data with the Roth et al. (1991) laboratory data, in which the proposer has more alternatives to choose from ($J_P = 201, J_R = 2$), the size of one proposer's strategy set is larger. In this case, as most offers were concentrated on even hundreds, I group the offers into 11 categories.[17]

Further, as pie size c and the coefficient λ of certitude are interchangeable in the logit response functions $\sigma_P(s)$ and $\sigma_R(s)$, the model also implies that increasing the size of the pie has the same effect as increasing the coefficient of certitude (even if the players' strategy sets have the same size across games). Inequality aversion b is not influenced by a variation of the pie size c. Therefore, its estimate is unaffected by the stakes of the game. These implications of the stakes, when comparing λ between games with different pie size, are discussed in detail in the paper within the later comparison of estimates across studies. Next, I derive the likelihood function for the UG to complete the presentation of the model.

4.3 Log-Likelihood function

The joint density of $k \in \{1, ..., K\}$ independent and identically distributed observations is given by multiplying the probabilities to observe each individual outcome (see, for instance, Greene, 2003, p. 468). As, according to Greene (2003, p. 468), "this joint density is the likelihood function," it is straightforward to derive the likelihood function for the UG. The difficulty outside the QRE framework is that typically there is no theory to impute the probability $\sigma_P(s)$ of observing a particular offer s, and probability $\sigma_R(s)$ of acceptance. This difficulty to obtain $\sigma_P(s)$ and $\sigma_R(s)$ notwithstanding, the general log-likelihood of observing a particular data sample of k observations in an UG is:[18]

[17] All findings remain unaltered in a cross check with the original data.

[18] See, for instance, also Yi (2005, pp. 325-26).

$$\ln L = \sum_k d_k \ln\left[\sigma_P(s) * \sigma_R(s)\right] + (1-d_k)\ln\left[\sigma_P(s) * (1-\sigma_R(s))\right]. \tag{5}$$

Note: d_k is a dummy for acceptance of s

With the logit response functions $\sigma_P(s)$ and $\sigma_R(s)$, which define the probabilities of each player choosing a particular strategy in the UG, it is possible to estimate decision error and the parameter of the social preference function by standard maximum likelihood estimation. The estimates of inequality aversion and decision error thereby obtained are consistent and asymptotically efficient. Their variances are given by the negative expectation of the information matrix. As the estimates are obtained maximizing the likelihood function of the UG, the likelihood ratio test, Wald test and Lagrange multiplier test can be used to test hypotheses on the model's parameters and the tests are asymptotically equivalent.[19] Consequently, in the logit equilibrium model it is possible to statistically test for the relevance of the social preference parameter b as well as decision error λ.

The identification the parameters b and λ is warranted through the functional form of the employed utility function. The coefficient of certitude λ is affected linearly and inequality aversion b non-linearly by payoff variations. Inequality concern b disappears only from the utility function under a rejection or when the equal split occurs. Any variation in responder behavior for the equal split could therefore be used to estimate the coefficient of certitude λ independently of functional form of inequality aversion. That is, the subset of the data, where $s = 0.5$, could in principal be used to test for a linear functional form of inequality aversion, as suggested by Fehr and Schmidt (1999), in the model. However, as no rejections are observed in the Zimbabwean dataset when the responders offer half, such a test is not pursued. Moreover, disregarding the full range of offers to test for the functional form of the utility function in the model yields conclusions based on a point estimate that could be rather unrepresentative for all observed play. Other robustness tests, which maintain the quadratic influence of inequality aversion, are reported later in the paper. At first, the next section presents the estimation results and goodness-of-fit tests of the above model on the UG data from the Zimbabwean field experiments.

[19] For discussion and proofs of these general maximum likelihood properties see, for instance, Greene (2003, pp. 470-492).

5 Estimation

5.1 Results and Predictive Success

While, for notational conciseness, during the model's introduction I used the coefficient of certitude λ, I henceforth present estimation results in terms of its inverse the decision error $\mu = \lambda^{-1}$. The estimation of the model as defined by equations (2) to (5) is at first on all 242 observations. In subsequent steps I estimate the model for female and male players, and then for resettled and traditional villagers allowing for various forms of heterogeneity.[20] I test all hypotheses about players' heterogeneity with likelihood ratio tests (henceforth LR-test). The LR-test statistic $LR = -2\ln(L_R/L_U)$ for a hypothesis, where L_R and L_U are the likelihood function evaluated at the parameter estimates of the restricted and the unrestricted model, is easy to compute after estimating the restricted and the unrestricted model. It is asymptotically distributed according to a chi-square distribution with the number of degrees of freedom equal to the number of restrictions being tested.[21] The likelihood-ratio test rejects the null hypothesis if the value of the statistic exceeds the appropriate critical value from the chi-squared tables.

Formally, I consider first the null hypothesis $H_2 : b = \mu = 0$ that the value of all parameters is equal to zero against the alternative $H_1 : b \neq 0, \mu \neq 0$; i.e. the hypothesis that a model of pure self-interest and without decision error is sufficient to explain the data. Then I test the individual significance of decision error $H_3 : \mu = 0 | b \neq 0$ and inequality aversion $H_4 : b = 0 | \mu \neq 0$, given the untested parameter's inclusion in the model, against the same alternative H_1. Table 2 summarizes the estimation results, the log-likelihood of the models, and the value of the LR-test statistics for the three aforementioned tests:

[20] This is, including interaction dummies for gender/resettlement status with inequality aversion and decision error in the estimation.

[21] Tabulated chi-square values: $\chi(0.01, 1) = 6.64$, $\chi(0.05, 1) = 3.84$, $\chi(0.1, 1) = 2.70$. Further, $\chi(0.01, 2) = 9.21$, $\chi(0.05, 2) = 5.99$, $\chi(0.1, 2) = 4.61$. Significance level and degrees of freedom are in parentheses.

Table 2: *Decision Error Symmetric Social Preferences*[22]

hypothesis	sample	N	μ	b	$-\ln(L)$	LR
$H_1 : \mu \neq 0, b \neq 0$	all	242	8.217	9.377	442.294	alternative
			(0.405)	(0.981)		hypothesis
$H_2 : \mu = b = 0$	all	242	0	0	∞	∞^{***}
			$(-)$	$(-)$		
$H_3 : \mu = 0, b \neq 0$	all	242	0	b	∞	∞^{***}
			$(-)$	$(-)$		
$H_4 : \mu \neq 0, b = 0$	all	242	9.757	0	591.268	297.948***
			(2.997)	$(-)$		

Note: Standard errors are in parentheses.

***/**/*Indicates rejection of the restriction (null) at the 1/5/10 percent significant level.

All parameters have the expected sign and are, jointly and individually, significant at the 1 percent level. The mean offer in the sample is 44.01 percent of the pie and mean acceptance in the sample is 92.15 percent. The model predictions for these two moments are 43.99 percent and 78.95 percent, respectively. The correlations between observed behavior $\sigma_i(s)$ and predictions $\widehat{\sigma}_i(s)$, for $i \in \{P, R\}$, are high, with a Pearson correlation coefficients of $\rho_P = 0.752$ and $\rho_R = 0.536$. The model fit is further illustrated in table 3 and figure 1, which compare observed versus predicted play. Figure 2, in contrast, graphs the standard game theoretic prediction against the observed play.

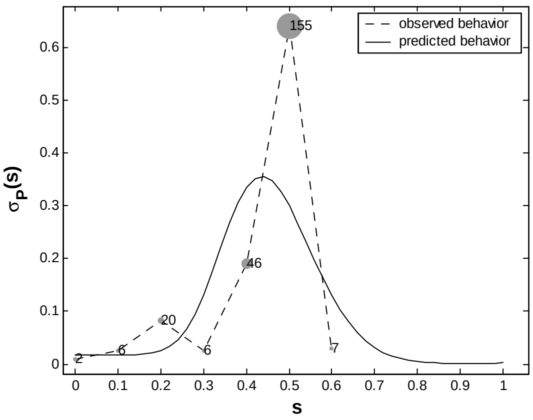
[22] The deterministic models cannot be estimatied. The likelihood function takes value minus infinity as the deterministic model fits the data imperfectly.

Table 3: *Observations vs. Model Prediction*

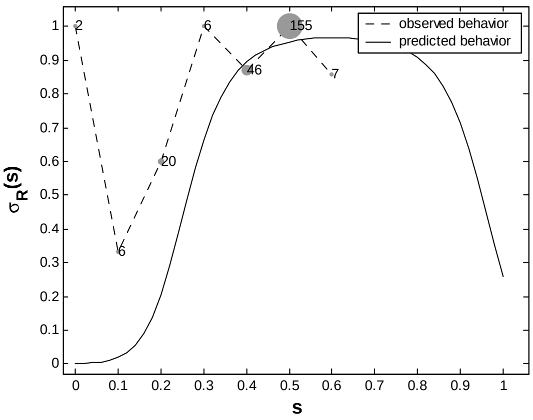
offer	observation		prediction	observation		prediction
s	N	$\sigma_P(s)$	$\widehat{\sigma}_P(s)$	N	$\sigma_R(s)$	$\widehat{\sigma}_R(s)$
0	(2)	0.008	0.017	(2/2)	1	0.001
0.1	(6)	0.025	0.017	(2/6)	0.333	0.019
0.2	(20)	0.083	0.027	(12/20)	0.600	0.206
0.3	(6)	0.025	0.131	(6/6)	1	0.665
0.4	(46)	0.190	0.336	(40/46)	0.870	0.896
0.5	(155)	0.641	0.301	(155/155)	1	0.955
0.6	(7)	0.029	0.132	(6/7)	0.857	0.967
0.7	(0)	0	0.032			0.958
0.8	(0)	0	0.005			0.909
0.9	(0)	0	0.001			0.713
1	(0)	0	0.003			0.260
	242	$\rho_P = 0.752$		242	$\rho_R = 0.536$	

Figure 1: *Symmetric Inequality Aversion and Decision Error*

$\mu = 8.217, b = 9.377$



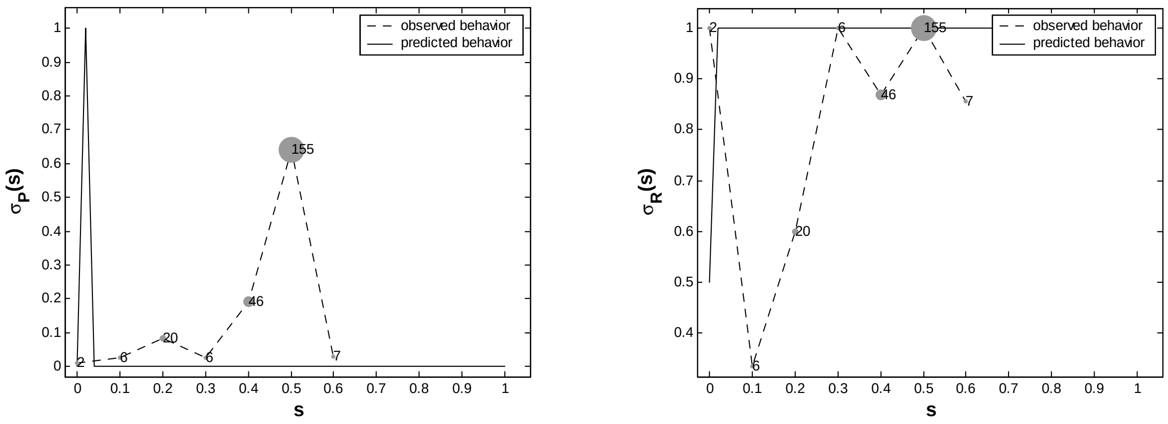
Proposer‘s probability to make an offer



Responders‘s probability to accept an offer

Figure 2: *Pure Self-Interest*

$\mu = b = 0$



Proposer‘s probability to make an offer — Responders‘s probability to accept an offer

As gender information about the UG participants was collected, and the games were played in two different types of villages, I next investigate if one set of parameter estimates is sufficient to explain behavior across the subject groups identified by these two characteristics. The previous estimation of the unrestricted model on the full dataset, as reported in table 2, is now equivalent to imposing the restriction that the parameter values are the same across gender and village type. In the tests of systematic differences amongst distinct participant groups this model becomes an important null hypothesis. More precisely, let subscripts f, m and r, t represent the estimates for the sample split by gender and resettlement status, respectively. The null hypothesis H_1, that one set of parameters (b, μ) is sufficient to explain all behavior, is tested against the alternatives $H_5 : b_f \neq b_m, \mu_f \neq \mu_m$, $H_6 : b_f \neq b_m | \mu_f = \mu_m$ and $H_7 : \mu_f \neq \mu_m | b_f = b_m$, that inequality aversion and decision error are jointly or individually different between the sexes, as well as $H_9 : b_f \neq b_m, \mu_f \neq \mu_m$, $H_{10} : b_r \neq b|, \mu_r = \mu_t$ and $H_{11} : \mu_r \neq \mu_t | b_r = b_t$, that inequality aversion and decision error differ jointly or individually by resettlement status. Tables 4 summarize the additional estimation results and findings from these and further hypothesis tests.

Table 4.1: *Decision Error and Symmetric Social Preferences in Subsamples*

hypothesis		sample	N	μ	b	$-\ln(L)$
$H_5:$	$\mu_f \neq \mu_m$	female	111	7.912 (0.578)	9.102 (1.484)	198.765
	$b_f \neq b_m$	male	131	8.466 (0.570)	9.567 (1.361)	243.285
$H_6:$	$\mu_f = \mu_m$	female	111	8.216 (0.956)	9.230 (1.431)	198.897
	$b_f \neq b_m$	male	131	8.216 (0.656)	9.498 (1.193)	243.388
$H_7:$	$\mu_f \neq \mu_m$	female	111	7.929 (0.578)	9.362 (1.252)	198.780
	$b_f = b_m$	male	131	8.456 (0.556)	9.362 (1.165)	243.297
$H_8:$	$\mu_f \neq \mu_m$	female	111	9.598 (3.471)	0 (−)	270.351
	$b = 0$	male	131	9.894 (3.370)	0 (−)	320.878
$H_9:$	$\mu_r \neq \mu_t$	resettled	180	7.597 (0.423)	12.638 (1.668)	303.493
	$b_r \neq b_t$	traditional	62	8.708 (0.962)	4.394 (1.279)	129.364
$H_{10}:$	$\mu_r = \mu_t$	resettled	180	7.844 (0.539)	12.631 (1.762)	303.655
	$b_r \neq b_t$	traditional	62	7.844 (1.108)	3.911 (1.186)	129.870
$H_{11}:$	$\mu_r \neq \mu_t$	resettled	180	7.605 (0.376)	9.805 (1.015)	305.108
	$b_r = b_t$	traditional	62	10.164 (1.030)	9.805 (1.584)	133.861
$H_{12}:$	$\mu_r \neq \mu_t$	resettled	180	9.918 (3.319)	0 (−)	440.973
	$b = 0$	traditional	62	9.275 (1.452)	0 (−)	150.149

Note: Standard errors are in parentheses.

Table 4.2: *Hypotheses tests in Subsamples*

			alternative hypothesis								
			all	female vs. male				resettled vs. traditional			
null hypothesis			H_1	H_5	H_6	H_7	H_8	H_9	H_{10}	H_{11}	H_{12}
H_1 :	μ	b	-	0.490	0.019	0.435	-	18.875***	17.540***	6.652***	-
H_2 :	$\mu = 0$	$b = 0$	∞***	∞***	∞***	∞***	∞***	∞***	∞***	∞***	∞***
H_3 :	$\mu = 0$	b	∞***	∞***	∞***	∞***	∞***	∞***	∞***	∞***	∞***
H_4 :	μ	$b = 0$	298***	298***	298***	298***	0.080	317***	316***	305***	0.292
H_5 :	$\mu_f \neq \mu_m$	$b_f \neq b_m$	-	-	-	-	-	-	-	-	-
H_6 :	μ	$b_f \neq b_m$	-	0.470	-	-	-	-	-	-	-
H_7 :	$\mu_f \neq \mu_m$	b	-	0.054	-	-	-	-	-	-	-
H_8 :	$\mu_f \neq \mu_m$	$b = 0$	-	298***	-	298***	-	-	-	-	-
H_9 :	$\mu_r \neq \mu_t$	$b_r \neq b_t$	-	-	-	-	-	-	-	-	-
H_{10} :	μ	$b_r \neq b_t$	-	-	-	-	-	1.335	-	-	-
H_{11} :	$\mu_r \neq \mu_t$	b	-	-	-	-	-	12.222***	-	-	-
H_{12} :	$\mu_r \neq \mu_t$	$b = 0$	-	-	-	-	-	317***	-	304***	-

Note: ***/**/*Indicates rejection of the restriction (null) at the 1/5/10 percent significant level.

A dash indicates that an *LR*-Test is not feasible as the null is no nested hypothesis within the alternative.

To obtain the likelihood function value of the partially or unrestricted models, I add the likelihood function values from the estimation of the subsamples under the according hypothesis. For instance, if testing for a joint gender difference, i.e. H_1 against H_5, the values $\ln(L_{H_1}) = -442.294$, from table 2, and $\ln(L_{H_5}) = -(198.765 + 243.285)$, from table 4.1, yield the reported statistic $LR = -2\ln(L_{H_1}/L_{H_5}) = 0.490$.

Overall, the structural analysis of gender and resettlement status confirms the behavioral differences suggested by the descriptive statistics of the UG in table 1. The hypothesis H_1 of unique inequality aversion and same decision error cannot be rejected in favor of any sex specific preferences at the 1, 5 or 10 percent level. Explicitly, there is no evidence for gender differences in decision error and inequality aversion jointly (H_5) or individually (H_6, H_7). Pure decision error models, without (H_4) and with gender specific error (H_8), are rejected in favor of the models including gender specific (H_5, H_6) or unique (H_7) inequality aversion at the 1 percent level. Moreover, also when ex ante excluding inequality aversion from the model (H_4) and the alternative specification (H_8) equal preferences across gender cannot be rejected. Hence, males err to the same extend as their female counterparts when making a decision in the UG and both sexes also share a unique concern for inequality aversion.

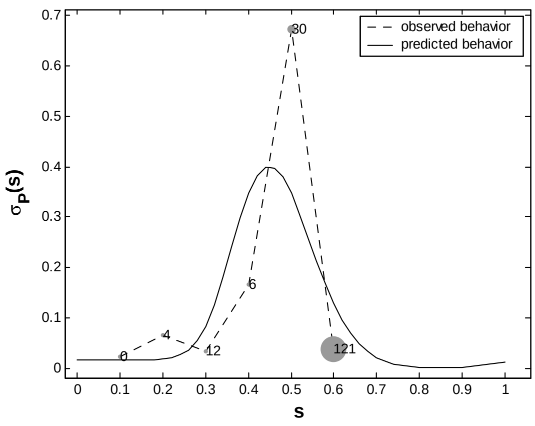
This result is different from Saad and Gill (2001), who, however, inform payers about their opponent's sex. Solnick (2001) also investigates explicitly gender differences in the UG in three distinct treatment conditions. In her baseline, players were labeled by number and thus did not know each other's gender. Under the treatment conditions, exchange of first names, should convey gender information, to disguise that gender is the issue of interest. She finds that offers in the treatment condition are affected by gender, with men attracting higher offers, particularly from female players, and players from both genders choose a higher minimum acceptable offer when facing a female player. In contrast, and in line with my findings, no gender pattern emerges in her data, when the co-players sex was unknown. According to Solnick (2001), overall only "few [further] UG studies include findings on gender, though gender is not the main variable of interest," in which neither player knew the gender of her/his partner. Murnighan and Saxon (1998) find in an investigation of an UG played by children that girls (from kindergarten to 9th grade) tend to make more generous offers than boys. Finally, Solim and Roth (1998) report from a high-stake UG played in Slovakia that men were less likely to reject offers. Such a further distinction of gender by player role lies outside the aim of this paper. Overall, the role of gender under anonymity appears minor or insignificant.[23] This paper adds evidence to the latter.

[23] Various laboratory experiments have examined the effect of gender in bargaining and strategic behavior. Overall, there are mixed results. See, for instance, Ball and Cech (1996).

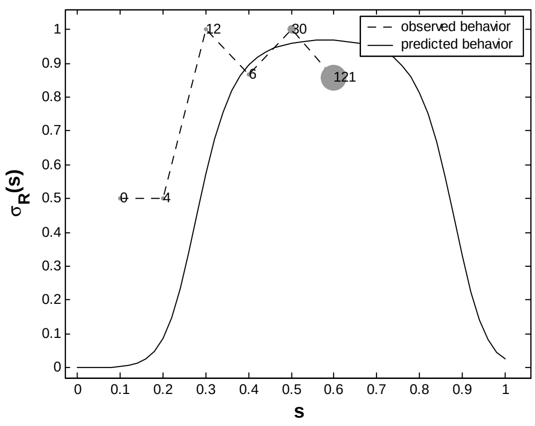
Comparing the unrestricted and various restricted models for resettled and traditional villages reveals that homogeneity of villagers (H_1) is rejected in favor of all models with different decision error and/or different inequality aversion (H_9, H_{10}, H_{11}) at the 1 percent level. Moreover, the hypothesis of heterogeneity in inequality aversion only (H_{10}) cannot be rejected against heterogeneity in inequality aversion and decision error (H_9) at the 1, 5 or 10 percent level, while H_{11}, i.e. heterogeneity in decision error only, is rejected against H_9 at the 1 percent level. That is, there is strong evidence for heterogeneity in inequality aversion across villager types but no reason to believe that resettled and traditional villagers make distinct decision errors in addition to their social preference difference. The unique estimate of decision error is 7.844 and inequality aversion, with 12.631 (resettled) and 3.911 (traditional), in resettled villages is estimated three times as high. Figures 3 show how these differences in inequality aversion causes predicted proposer and responder behavior to differ by resettlement status.

Figure 3.1: *Symmetric Inequality Aversion and Decision Error*

Resettled Villages $\mu_r = 7.844, b_r = 12.631$



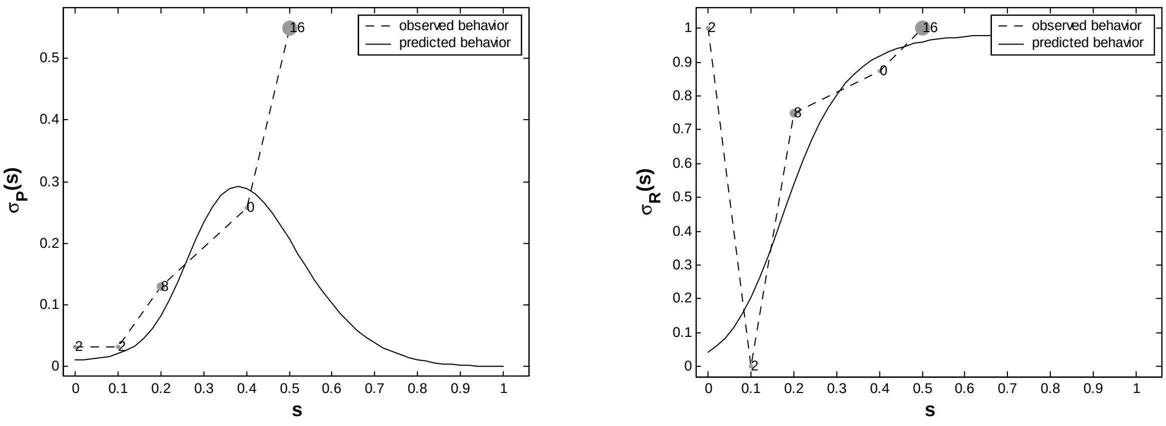
Proposer's probability to make an offer



Responders's probability to accept an offer

Figure 3.2: *Symmetric Inequality Aversion and Decision Error*

Traditional Villages $\mu_t = 7.844, b_t = 3.912$



Proposer's probability to make an offer

Responders's probability to accept an offer

Barr (2003a) studies Trust Game behavior in these villages to analyze the social consequences of the Zimbabwean land reform. Thereby she controls for external factors and village fixed effects. She (p. 614) concludes about the motives behind the trust behavior that: "Both the data and certain stylized facts suggest that altruistic motivations matter less and motivations relating to a desire to 'community-built' matter more in resettled communities." The higher value of inequality aversion in resettled villages and similar decision error in the resettled and traditional villages, are consistent with this desire to community build, as on the one hand inequality aversion supports sanctioning of non-social behavior in the Ultimatum Game, or captures an intrinsic motivation to cooperation in Public Good Games, while on the other hand, and in contrast to altruism, it can be overruled by self-interest in competitive situations like the Market Game.[24]

Barr (2004a) jointly investigates UG and Trust Game behavior in the Zimbabwean villages. She (p. 327) conjectures: "Rules about sharing and fairness do not appear to vary [between resettled and traditional villages]. The lower trust among the unrelated, resettled villagers appears to be due to a reduced ability to predict their fellow villagers' behavior, that is, they are less familiar with one another's behavioral characteristics than the more inter-related non-resettled villagers." The preference difference amongst the resettled and traditional villagers identified in this paper, while in line with Barr (2003a), is antipodal to

[24] For a discussion of these games and the according ERC predictions see Bolton and Ockenfels (2000).

the latter explanation because heterogeneous inequality aversion represents variation in fairness and equal decision error suggests similar unspecific heterogeneity and hence familiarity. This notwithstanding, the paper at hand complements Barr (2004a, 2003a) by offering an approach that directly brings the observed play to the preference level.

In a more general context than land reform, Henrich et al. (2005) analyze the impact of the cultural and economic environment on the formation of preferences across countries. They find that the variation in payoffs to cooperation (i.e., how important and how large is a group's payoff from cooperation in economic production?) and market integration (i.e., how much do people rely on market exchange in their daily lives?) across countries are good predictors for the observed differences in UG behavior. "In a regression, both PC [payoffs to cooperation] and MI [market integration] were highly significant, their (positive) normalized regression coefficients were large in magnitude (about 0.3), and the two measures jointly explained 68 percent of the variance." (Henrich et al. 2005, p. 76) This finding supports the hypothesis of stronger social preferences in resettled villages due to higher returns to cooperation.

Overall, given the direction of the difference in inequality aversion between resettled and traditional villagers, land reform itself could have had a direct impact on the formation of preferences because people in resettled areas came together from different traditional villages to commonly make a new start. However, for a causal identification of an impact of land reform on the found preference difference, I would need to control for selection into the land reform project. This lies outside the scope of this paper. Nonetheless, the model applied to field experimental data establishes a new clearly interpretable, stylized fact about the not directly observable incentive structure in two distinct environments, which is consistent with ideas about the formation of social preferences put forward by others: Resettled and traditional villagers differ systematically in exerted inequality aversion.

5.2 Alternative Model Specifications and Robustness

The ERC social preference function (2) postulates average income in the comparison group as the social reference point. Any deviation of the own payoff from this reference point diminishes utility in a quadratic way. Other distributional models, as, for instance, the one suggested by Fehr and Schmidt (1999), assume that subjects dislike payoff difference to any other individual. The sum of these differences then lowers utility linearly. Both of the two models share the assumption that deviation from the reference point towards an advantageous financial situation matter as well as a disadvantageous situation. Hence, subjects are assumed to dislike being worse off and being better off. Fehr and Schmidt (1999), in addition, introduce two weights for the different direction of deviations from the reference point and thereby allow for a different utility weight of disadvantageous and advantageous relative utility. In applications of their model, advantageous utility is also sometimes assumed not to matter. While this truncation of the Fehr and Schmidt (1999) model in part is undertaken for simplification, Bolton's (1991) comparative bargaining model directly employs asymmetric inequality aversion, and thereby explains various departures from purely self-interested behavior in bargaining experiments.

Since there is mixed evidence on the exact nature of inequality aversion, I test whether a symmetric or asymmetric specification of quadratic inequality aversion in the utility function is more appropriate for the given data. The characterization symmetric and asymmetric refers to whether the player dislikes being better of as well as being worse off or only being worse off. Formally, I test the hypothesis of $H_1 : b = 0 \forall \omega_i \geq 0.5$ against the alternative $H_2 : b \neq 0$, given $\mu \neq 0$, where under H_1 the ERC utility function (2) is restricted to:

$$u_i(x_i, x_j) = \begin{cases} c\left(\omega_i - \frac{b}{2}\left(\omega_i - \frac{1}{2}\right)^2\right) & \text{if} \quad \omega_i < \frac{1}{2} \\ c\omega_i & \text{if} \quad \omega_i \geq \frac{1}{2} \\ 0 & \text{if} \quad x_i, x_j = 0 \end{cases} \tag{6}$$

Both specifications H_1 and H_2 have intuitive appeal, but predict to some extent distinct behavior. To illustrate the difference consider the UG. Subjects with asymmetric social preferences H_1 (6) only care about not being behind. Proposers merely make positive offers, because they anticipate rejection of small ones. In turn, if subjects have symmetric social preferences H_2 (2) and their inequality aversion is high enough then proposers have an intrinsic motivation to share up to one-half of the surplus even if no rejections are expected. By similar reasons, responders with symmetric social preferences stop accepting offers at some point, as they become larger afterward an equal split was reached.

Often asymmetric inequality aversion is sufficient to explain as strategic behavior, what would seem irrational in the light of self-interest. De Bruyn and Bolton (2004) also test both ERC specifications and favor asymmetric inequality aversion specification to predict behavior across several bargaining games in laboratory experiments. However, as Henrich et al. (2005), Henrich (2000), and other anthropologically motivated studies since then show, there is great cultural diversity in behavior, including frequent rejection of offers bigger than half of the pie in Ultimatum Games in some societies. Frequency of this behavior is evidence for symmetric inequality aversion and is typically not observed in laboratory experiments in industrialized countries. A reported reason as for why rejection of high offers have manifested itself in some societies, in which economic success depends extensively on cooperation and sharing, is the fear from punishment after a potential inability to return a generous favor. On the social context which stipulates offers greater than 50 percent see also Hill and Gurven (2004), who study the UG behavior of Ache Indians of Paraguay.

The mixed evidence on the nature of inequality concern notwithstanding, the structure of the UG does not imply a particular choice between the models of symmetric and asymmetric inequality aversion. To my knowledge there is no reason to ex ante favor one specification over the other in the Zimbabwean context. Consequently, I apply both specifications to the data to test for the more appropriate choice. Table 5 summarizes the estimation results for the three possible models (including pure self-interest) that can be fitted in the quantal response framework. Each model is complemented by its deterministic counterparts without decision error:

Table 5: *Alternative Model Specifications*

		(A) self-interest		(B) symmetric		(C) asymmetric	
	obs.	det.	prob.	det.	prob.	det.	prob.
hypothesis		–	H_0	–	H_1	–	H_2
$-\ln L$		∞	591.267	∞	442.294	∞	468.000
μ		0	9.757	0	8.217	0	9.252
b		0	0	76.800	9.377	55.154	10.703
$\overline{s}$	0.44	0.100	0.310	0.500	0.440	0.400	0.417
$\varnothing$ Pr(Y)	0.92	1	0.880	0.714	0.790	0.714	0.770
ρ_P		-0.115	0.173	0.955	0.752	0.170	0.642
ρ_R		-0.333	0.315	0.370	0.536	0.370	0.531

Note: $H_0 : b = \mu = 0$, $H_1 : b = 0 \forall \omega_i \geq 0.5 | \mu \neq 0$, $H_2 : b \neq 0 | \mu \neq 0$

Deterministic models are fitted by calibration maximizing $\Sigma_{i \in \{A,B\}} \rho_i$.

Calibration allows for different b-values without changing the result.

In the standard equilibrium (A) proposer and responder are only motivated by monetary gains. Hence, the proposer offers the smallest share zero (or one unit) and the responder accepts. The model (B), in which both players exhibit symmetric inequality aversion is used in the paper. The last columns present the estimation results from a model (C) with asymmetric social preferences. The deterministic counterparts of all three specifications cannot be estimated but are calibrated with exemplifying b-values, pursuing the highest possible correlation between observed and predicted play. As the deterministic models imply the choice probabilities to be either zero or one, the likelihood function takes constant value minus infinity except for a perfect fit of the data. That is, until there is no variation in the data, decision error μ, which makes the model stochastic, is by definition significant, regardless of the deterministic utility specification.

Comparing the probabilistic models (A) with (B) and (C), in which $\mu \neq 0$, using a likelihood ratio test yields test statistics of $LR = 297.948$ for testing model (A) against model (B), i.e. H_0 against H_2, and $LR = 246.534$ for testing model (A) against model (C), i.e. H_0 against H_2, respectively. These values imply at the 1 percent level that inequality aversion b is an individually significant variable in a decision error model. When comparing the two inequality aversion specifications (C) and (B), i.e. when testing H_2 against H_1, the test statistic $LR = 51.402$ indicates further significance of advantageous inequality concern at the 1 percent level. In other words, an unrestricted symmetric model, as has been used above, is the appropriate model choice for the considered data. The correlation coefficients ρ_P (ρ_R) between the observed and predicted offer (acceptance) frequencies emphasize this point, because they indicate a better fit of model (B), in particular for the proposer's behavior.

It is noteworthy that, when judging the performance of the different possible models on the basis of the correlation coefficients, a deterministic-symmetric specification performs well too. This is due to the high concentration of offers of an equal split (155/242), which are all accepted in the sample. However, as the variance of offers becomes larger than in the Zimbabwean UG data, a deterministic model increasingly fails to capture a broader dispersion of offers.

6 Experimental Field vs. Laboratory Data

Existing lab-experimental data and its analyses have shown that people from many parts of the world (Europe, Asia, and North America) behave quite similarly in the UG. In a study from Jerusalem (Israel), Ljubljana (Slovenia), Pittsburgh (US) and Tokyo (Japan), Roth et al. (1991) find differences in bargaining behavior across countries, which they tentatively attribute to cultural differences, but note that a robust pattern emerges. Overall, proposers make similar mean offers, and responders frequently reject low, inequitable offers. This common pattern combined with some cross-country variation suggests the Roth et al. (1991) data as a benchmark for the comparison of results. Roth et al. invited subjects for a Market Game and an UG in the aforementioned four countries. Games were repeated ten times and each time subjects were re-paired randomly and anonymously. The stake in the Ultimatum Games was US$ 10, or the local currency equivalent, in all but one Pittsburgh treatment in which US$ 30 had to be divided.

I reestimate the noisy inequality aversion model for their small stake data using only the first round of bargaining to exclude learning effects as the game proceeds. After dropping 9 out of 10 rounds, 125 observations from ultimatum bargaining with the default stake remain. Table 6 presents the estimation results for this data by country:

Table 6: *Decision Error and Symmetric Social Preferences in Roth et al. (1991) Data grouped in Bins of 100 Token*[25]

	N	μ	$\mu_{adjusted}$	b	$-\ln(L)$
all	125	10.827	1.083	10.338	286.566
Jerusalem	30	11.356	1.136	7.307	69.877
Tokyo	29	14.203	1.420	9.170	75.786
Ljubljana	30	9.388	0.939	14.364	62.294
Pittsburgh	36	9.361	0.936	12.136	75.035

The Zimbabwean estimate for inequality aversion $b = 9.377$ lies in the medium range of the US$ 10 games, which have (non-standardized) b estimates ranging from 7.307 in Jerusalem to 14.364 in Ljubljana. The comparison of decision error is problematic and needs further explanation. With the way the model is constructed, pie size influences the estimation results for decision error $\mu = 1/\lambda$. Substituting social preferences (2), for instance, into the responder's acceptance probability (3) yields:

$$\sigma_R(s) = \frac{e^{\lambda c\left(\omega_i - \frac{b}{2}\left(\omega_i - \frac{1}{2}\right)^2\right)}}{1 + e^{\lambda c\left(\omega_i - \frac{b}{2}\left(\omega_i - \frac{1}{2}\right)^2\right)}},$$

[25] A characteristic of the Roth et al. (1991) data is that games were played for tokens. In the Ultimatum Game 1000 token, corresponding to US$ 10, could be divided in units of 5. The table shows estimation results, when this data is grouped in bins of width 100 token. Grouping corresponds well with actual behavior, where offers are concentrated on multiples of 100 token. This approach is also taken by Costa-Gomes and Zauner (1998). The appendix A5 provides a comparison to the estimation with the original data. Notwithstanding whether grouping is the right approach to this data, it yields decision error estimates of approximately one-half of the estimates from the raw data. With this difference constant and low the subsequent interpretation and comparisons remain qualitatively unaffected by the grouping.

i.e. $\lambda c = c/\mu = K$ and therefore the decision error estimate depends directly on the pie size c. Since within the Roth et al. (1991) data bargaining in each country is for a local currency equivalent of US$ 10, the decision error estimates are comparable.[26]

Taking into account that the per capita GDP in the US at the time of the games 1990 was US$ 26,141 as compared to the per capita GDP in Zimbabwe of Zim$ 1,986 in the year 2000 (World Values Survey), could be an approximation for the 'true' value of the relative stakes, i.e. in comparable units the bargaining stake in Zimbabwe was about $(Zim\$ \ 50 \ / \ Zim\$ \ 1,986) \ / \ (US\$ \ 10 \ / \ US\$ \ 26,141) \approx 65$ times the stake of the US bargainers and possibly the others. As within the cross country comparison of Roth et al. (1991) the pie size actually specified in the estimation routine does not matter (every game had US$ 10 stakes in the local currency equivalent), the error estimates μ tabulated in table 5 stem from an estimation in which the Zimbabwean pie size was used as the numeraire. Conservatively adjusting these estimates for the lower per capita GDP in Zimbabwe by dividing them by 10 (instead of 65) already indicates that in the Zimbabwean data, for which $\mu = 8.217 >> \mu_{adjusted}$, decision error in behavior is much more prevalent. Decision error, in turn, is part of the model and does not imply a bad fit of the data.

A second interesting benchmark of comparison for the Zimbabwean findings is Henrich et al. (2005), who compare UG studies in 15 small-scale economies. A subset of the Zimbabwean data at hand has been part of this study. It shows that there is considerably more cross cultural variation than found by Roth et al. (1991). Under the aforementioned caveat of the non-comparability of the decision error estimate across games with different pie size, I reestimate the model on further countries of the Henrich et al. (2005) data. Because the smallest unit and hence the number of choices for the proposer varied between games, which in turn adds or takes away artificial noise due to the discrete specification of the proposer's probability $\sigma_P(s)$ to make an offer s, I consider only 4 of 10 more available countries from this dataset. The four additional countries chosen - Tanzania, Papua New Guinea, Paraguay and Kenya - had an UG setup similar to the Zimbabwean UG ($J_P = 10, J_R = 2$), in which offers could be increased in 10 percent increments of the total. Table 7 reports the estimation results for these countries:

[26] Arguable if further assuming that the local currency equivalent reflects similar purchasing power, such that participants in all countries enjoy the same utility from a US$ 1 payoff.

Table 7: *Decision Error and Symmetric Social Preferences in Henrich et al. (2005)*

	N	μ	b	$-\ln(L)$
all	217	8.762	3.881	482.320
Tanzania	55	7.764	1.116	129.639
PNG	55	12.722	4.909	145.414
Paraguay	51	6.940	10.078	89.410
Kenya	56	5.281	7.291	83.315

Note: PNG = Papua New Guinea

In comparison with the other field data, the Zimbabwean estimates for inequality aversion $b = 9.377$ lies in the upper range of the estimates. Decision error within the different field datasets is not directly comparable. However, the stakes in all Henrich et al. (2005) countries were between half a day's and a day's casual labor wage. Thus, when comparing the decision error estimates from the field data to the with factor 1/10 adjusted Roth et al. (1991) error estimates, it is evident, that the field data shares the common feature that decision error is substantially higher than in the data of Roth et al. (1991) from industrialized populations.

7 Conclusion

By applying a model of inequality aversion and noisy behavior to Ultimatum Game data from a field experiment, I find that neither self-interest nor decision error on their own are sufficient to explain observed bargaining behavior. A social preference function of inequality aversion, suggested by Bolton and Ockenfels (2000), embedded in a quantal response model allows for error in decision making and fits the data well. Overall, predicted behavior corresponds with 75 percent (54 percent) of the observed proposer (responder) behavior. Both parameters of the model, decision error μ (or equivalently its inverse the coefficient of certitude λ) and inequality aversion b are significant at the 1 percent level. The highest predictive success was achieved with a symmetric model of inequality aversion, implying that the Zimbabwean villagers dislike being better of as well being worse off than average.

Gender differences in preferences were not observed. When comparing resettled villages and traditional villages, an on average higher offer in resettled villages is attributed to a difference in inequality aversion, which in resettled villages is about triple of the estimate from the traditional villages. A 13 percent lower decision error in resettled villages is not significant, despite the fact of higher ethnic diversity within them, after the heterogeneity in inequality concern is considered. People in resettled villages, who stem from diverse traditional villages in the observed regions, made a new start after the Zimbabwean land reform in the early 1980s. This could suggest that a need for more cooperation in the enterprise to start a new community has entered peoples preference over this time; i.e. as a self-fulfilling prophecy reinforcement of a social norm to cooperate may manifest in people's preferences and thereby become a credible reinforcement device for cooperation like inequality aversion.

As the use of experimental economic field data from non-student populations and/or from a wide cultural background is a recent development and the application of a quantal response model to such data is new, I reestimated the model with other datasets from Roth et al. (1991) and Henrich et al. (2005). The comparison indicates more decision error in non-industrialized than in industrialized populations. Decision error, which is an essential part of the model, thereby does not suggest a bad fit of the data, but rather that players err more in the way captured when making their decisions. The result of noisier behavior holds for the Zimbabwean analysis as well as when considering other small-scale societies in non-industrialized countries. This suggests that higher decision error contains a component, which is a commonality to field experiments. The latter is striking as the stakes in the small-scale societies are substantially larger than in the industrialized comparison groups and thus careful decision making is tied to higher benefits. If decision error is interpreted as a shortcut to the unobserved difference between players, greater heterogeneity amongst field players could be a reason for the behavioral difference. If one believes more in a literal interpretation of the decision error, higher error may suggest that Ultimatum Game bargaining situations are closer to our everyday environment.

Acknowledgements *I am grateful to Abigail Barr, Alvin Roth, Joe Henrich and the co-authors of the two latter for sharing their experimental data. Naturally, my thank you extends to the experimentees, in particular the villagers from Zimbabwe, for participating in the adventure of carrying economic experiments into the field. I wish to thank Abigail Barr, Pascal Courty and Karl Schlag for stimulating discussions and Arnaud De Bruyn helpful comments. The paper benefitted also from discussions at the EUI and at the International Workshop on Behavioural Game Theory and Experiments in Capua 2006. Part of this paper was written under the supervision of Abigail Barr while I was a visiting student of the Centre for the Study of African Economies at the University of Oxford. Her guidance and* *the hospitality of the centre are gratefully acknowledged. Needless to say, all errors are my responsibility.*

References

[1] Anderson, S. P., Goeree J. K. and C. A. Holt, "A theoretical analysis of altruism and decision error in public goods games,"Journal of Public Economics 70, 1998, pp. 297-323.

[2] Ball, S. B. and P.-A. Cech, "Subject pool choice and treatment effects in economic laboratory research," in Research in experimental economics, Volume 6, edited by R. M. Isaac, Greenwich, CT: JAI Press, 1996, pp. 239-92.

[3] Barr, A., "Kinship, Familiarity, and Trust: An Experimental Investigation," in J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr and H. Gintis, Eds., Foundations of Human Sociality, Oxford University Press, Oxford, 2004a, pp. 305-334.

[4] Barr, A., "Forging effective new communities: The evolution of civil society in Zimbabwean resettlement villages," World Development, 32, 2004b, pp. 1753-66.

[5] Barr, A., "Trust and expected trustworthiness: Experimental evidence from Zimbabwean villages," The Economic Journal, 113, 2003a, pp. 614-30.

[6] Barr, A., "Risk pooling, commitment, and information: An experimental test of two fundamental assumptions," Working Paper No. 2003-05, The Centre for the Study of African Economies, University of Oxford, 2003b.

[7] Barr, A. , "From strangers to neighbors: Identifying the pre-conditions for social capital development," mimeo, The Centre for the Study of African Economies, University of Oxford, 2002.

[8] Barr, A., "Social dilemmas and shame-based sanctions: Experimental results from rural Zimbabwe," Working Paper No. 2001-11, The Centre for the Study of African Economies Working Paper Series, University of Oxford, 2001.

[9] Binmore, K., J. McCarthy, G. Ponti, L. Samuelson, and A. Shaked, "A backward induction experiment," Journal of Economic Theory, 104, 2002, pp. 48-88.

[10] Bolton, G. E., "A comparative model of bargaining: Theory and evidence," American Economic Review, 81, 1991, pp. 1096-1136.

[11] Bolton, G. E. and Axel Ockenfels, "ERC: A theory of equity, reciprocity, and competition," American Economic Review, 90, 2000, pp. 166-93.

[12] Camerer, C. F. and R. H. Thaler, "Anomalies: Ultimatums, dictators and manners," Journal of Economic Perspectives, 9, 1995, pp. 209-19.

[13] Charness, G. and M. Rabin, “Social preferences: Some simple tests and a new model,” Quarterly Journal of Economics, 117, Aug. 2002, pp. 817-69.

[14] Cooper, D., and J. van Huyck, "Evidence on the Equivalence of the Strategic and Extensive Form Representation of Games," Journal of Economic Theory, 110, 2003, pp. 290-308.

[15] Costa-Gomez, M. and K. G. Zauner, "Ultimatum bargaining behavior in Israel, Japan, Slovenia, and the United States: A social utility analysis," Games and Economic Behavior, 34, 2001, pp. 238-269.

[16] De Bruyn, A. and G. Bolton, "Predicting the role of fairness in bargaining - Out-of-sample estimates from a social utility model with Quantal Response," ISBM Working Paper No. 1-2004, Pennsylvania State University, 2004.

[17] Epps, T. W. and K. Singleton, “An Omnibus Test for the Two-Sample Problem Using the Empirical Characteristic Function,” Journal of Statistical Computation and Simulation, 26, 1986, pp. 177-203.

[18] Fehr, E., U. Fischbacher, B.v. Rosenbladt, J. Schupp and G. Wagner, "A nation-wide laboratory - Examining trust and trustworthiness by integrating experiments in representative surveys," Institute for Empirical Research in Economics, University of Zürich, Working Paper No. 141, 2003.

[19] Fehr, E. and K. M. Schmidt, “A theory of fairness, competition, and cooperation,” Quarterly Journal of Economics, 114, 1999, pp. 817-68.

[20] Gale, J., Binmore, K. and L. Samuelson, "Learning to be imperfect: The Ultimatum Game," Games and Economic Behavior, 8, 1995, pp. 56-90.

[21] Goeree, J. K., C. A. Holt and T. Palfrey, "Regular Quantal Response Equilibrium," Social Science Working Paper No. 1219, California Institute of Technology, 2005.

[22] Goeree, J. K. and C. A. Holt, "Asymmetric inequality aversion and noisy behavior in alternating-offer bargaining games," European Economic Review, 2000, 44, 4-6, pp. 1079-89.

[23] Greene, W. H., "Econometric analysis," 5th edition, Prentice Hall, Upper Saddle River, NJ, 2003.

[24] Henrich, J., "Does culture matter in economic behavior? Ultimatum Game bargaining among the Machiguenga of the Peruvian Amazon," American Economic Review, 90, 2000, pp. 973-79.

[25] Henrich, J., R. Boyd, S. Bowles, H. Gintis, E. Fehr, C. Camerer, R. McElreath, M. Gurven, K. Hill, A. Barr, J. Ensminger, D. Tracer, F. Marlow, J. Patton, M. Alvard, F. Gil-White and N. Henrich, "Economic Man' in Cross-Cultural Perspective: Ethnography and Experiments from 15 small-scale societies," Behavioral and Brain Sciences, 28, 2005, pp. 795-855. Previously published as Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis and R. McElreath, "In search of homo economicus: Behavioral experiments in 15 small-scale societies," American Economic Review, 91, May 2001, pp. 73-78.

[26] Hill, K. and M. Gurven, "Economic experiments to examine fairness and cooperation among the Ache Indians of Paraguay," in J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr and H. Gintis, Eds., Foundations of Human Sociality, Oxford University Press, Oxford, 2004a, pp. 382-412.

[27] Johnson, E., Camerer, C., Sen, C. and T. Rymon, "Detecting failures of backward induction: Monitoring information search in sequential bargaining," Journal of Economic Theory, 104, 2002, pp. 16-47.

[28] Kohler, S. ,"Fairness vs. Social Welfare in Experimental Decisions", Working Paper No. ECO2005/11, European University Institute, 2005.

[29] McKelvey, R., and T. Palfrey, "Quantal Response Equilibria for normal-form games," Games and Economic Behavior, 10, 1995, pp. 6-38.

[30] McKelvey, R., and T. Palfrey, "Quantal Response Equilibria for extensive-form games," Experimental Economics, 1, 1998, pp. 9-41.

[31] Murnighan, J. K. and M. S. Saxon, "Ultimatum bargaining by children and adults," Journal of Economic Psychology, 19, 1998, pp. 415-45.

[32] Myerson, R. B., "Refinements of the Nash Equilibrium concept," International Journal of Game Theory, 7, 1978, pp. 73-80.

[33] Offerman, T., A. Schram and J. Sonnemans, "Quantal Response models in step-level Public Good Games," European Journal of Political Economy, 14, 1998, pp. 89-100.

[34] Palfrey, T. R. and J. E. Prisbrey, "Anomalous behavior in public goods experiments: How much and why?," American Economic Review, 7, 1997 pp. 829-46.

[35] Prasnikar, V. and A. E. Roth, "Considerations of fairness and strategy: Experimental data from sequential games," Quarterly Journal of Economics, 107, 1992, pp. 865-88.

[36] Rabin, M., "Incorporating fairness into game theory and economics," American Economic Review, 83, 1993, pp. 1281-1302.

[37] Roth, A. E., "Bargaining experiments," in J. H. Kagel and A. E. Roth, Eds., Handbook of experimental economics, Princeton University Press, Princeton, NJ, 1995, pp. 253-348.

[38] Roth, A. E. and I. Erev, "Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term," Games and Economic Behavior, 8, 1993, pp. 164-212.

[39] Roth, A. E., V. Prasnikar, M. Okuno-Fujiwara and S. Zamir, "Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An experimental study," American Economic Review, 81, 1991, pp. 1068-95.

[40] Saad, G. and T. Gill, "Sex differences in the Ultimatum Game: An evolutionary psychology perspective," Journal of Bioeconomics, 3, 2001, pp. 171-93.

[41] Schotter, A., K. Weigelt, and C. Wilson, "A laboratory investigation of multiperson rationality and presentation effects," Games Economic Behavior, 6, 1994, pp. 445-468.

[42] Selten, R., "Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit," Zeitschrift für die Gesamte Staatswissenschaft, 121, 1965, pp. 301-24, pp. 667-89.

[43] Slonim, R. and A. E. Roth, "Learning in high stakes Ultimatum Games: An experiment in the Slovak Republic, " Econometrica, 66, 1998, pp. 569-96.

[44] Solnick, S., "Gender differences in the Ultimatum Game," Economic Inquiry, 39, 2001, pp. 189-200.

[45] World Values Survey, http://www.worldvaluessurvey.org/

[46] Yi, K. O., "Quantal-response equilibrium models of the ultimatum bargaining game," Games and Economic Behavior, 51, 2005, pp. 324-48.

8 Appendix A

Table A1: *Village Names, Types, Areas and Ultimatum Game Behavior*

	Village	Re-settled	Area	Vil. Code	N	proportion of females	mean offer	accep-tance
1	Chitepo	Yes	Mupfurudzi	11	13	0.39	0.49	0.85
2	Mudzinge	Yes	Mupfurudzi	12	14	0.36	0.43	0.87
3	Muringamombe	Yes	Mupfurudzi	13	11	0.36	0.5	1
4	Mupedzanhamo	Yes	Mupfurudzi	15	13	0.46	0.44	0.85
5	Zvataida	Yes	Mupfurudzi	16	13	0.46	0.52	0.92
6	Tongogara	Yes	Mupfurudzi	17	20	0.3	0.36	0.9
7	Gwetera	Yes	Mupfurudzi	18	15	0.33	0.49	0.87
8	Zvomanyanga	Yes	Mupfurudzi	19	12	0.42	0.38	1
9	Village 10	Yes	Mutanda	24	8	0.13	0.48	1
10	Village 11a and b	Yes	Mutanda	25	10	0.4	0.48	1
11	Mungo	Yes	Sengezi	31	10	0.4	0.48	1
12	Goto	Yes	Sengezi	32	15	0.53	0.43	0.87
13	Rundu	Yes	Sengezi	33	10	0.9	0.5	1
14	Mawire East	Yes	Sengezi	36	16	0.82	0.44	1
15	Sengenda	No	near. Sengezi	94	10	0.6	0.32	1
16	Chigwedere	No	near. Sengezi	95	8	0.63	0.45	0.75
17	Guzemuka	No	near. Mutanda	96	8	0.5	0.48	1
18	Madziwana	No	near. Mutanda	97	12	0.5	0.4	0.83
19	Chechera	No	near. Mupfurudzi	98	12	0.42	0.4	0.83
20	Paswavaviri	No	near. Mupfurudzi	99	12	0.33	0.42	1
	Total				242	0.46	0.44	0.92

Table A2: *Comparison of Household Demographics, Wealth, Within-Village Linkages, and Village Ethnic Composition in Resettled and Traditional Areas (2000)*

	traditional villages	resettled villages	range of village means in resettled areas	range of village means in traditional areas
non-religious memberships	1.52 (143)	4.09* (557)	1.00 - 7.50	0.78 - 4.14
household size	5.90 (143)	9.39* (394)	5.73 - 14.20	5.00 - 6.88
women	0.27 (143)	0.28 (394)	0.20 - 0.34	0.20 - 0.32
young	0.41 (143)	0.40 (394)	0.33 - 0.50	0.42 - 0.46
aged	0.11 (143)	0.08* (394)	0.03 - 0.12	0.06 - 0.20
livestock (1999)	7.71 (145)	13.59* (568)	7.29 - 25.15	6.35 - 7.47
marriage ties	0.91 (188)	0.81 (753)	0.06 - 1.87	0.53 – 1.86
extended family ties	0.55 (188)	5.76* (753)	0.00 - 1.56	1.42 – 10.81
nuclear family ties	1.91 (188)	0.30* (753)	0.00 - 0.68	0.82 - 2.83
some initial social capital	1.00 (245)	0.83* (723)	0.42 - 1.00	1.00 - 1.00
ethnic dominance	49.67 (6)	40.77 (22)	0.19 - 1.00	0.32 - 0.70
ethnic diversity	5.33 (6)	5.95 (22)	1.00 - 11.00	4.00 - 6.00
number of households	46.17 (6)	37.00* (22)	13.00 - 64.00	34.00 - 63.00

Note: All variables refer to year 2000 values if not specified otherwise.

Numbers of observations in parentheses.

$^{***/**/*}$Indicates significant difference at 1/5/10 percent level.

Table A3: *Non-structural Offer Analysis*

Variable	Model 1	Model 2	Model 3	Model 4
constant	16.370***	16.101***	15.990***	15.000***
offer Accepted	10.360***	10.187***	10.200***	10.200***
traditional Village	-3.224	-2.401	-1.135	-
female Proposer	-1.352*	-0.272	-	-
female Proposer from traditional Village	3.887***/j**	-	-	-
tillage Dummies	all***	all***	all***	all***
R^2-adj.	38.10	36.14	36.37	36.37
	Model 5	Model 6	Model 7	Model 8
constant	25.115***	24.615***	14.100***	13.798***
offer Accepted	-	-	9.667***	9.469***
traditional Village	-3.800	-	-3.626***	-2.028***
female Proposer	-1.298	-	-1.066	-
female Proposer from traditional Village	3.194**	-	3.386**/j***	-
village Dummies	all***	all***	-	-
R^2-adj.	14.48	13.63	23.23	22.17

Note: 242 observations.

$^{***/**/*}$Indicates 1/5/10 percent significance level. j-/ j* indicate joint in/significance of interaction variable and individual interaction terms.

Table A4: *Non-structural Acceptance Analysis*

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
constant	0.167*	0.204**	0.203***	0.846***	0.410***
offer	0.027***	0.026***	0.026***		0.022***
traditional Village	0.047	0.177*	0.177*	0.154	0.095*
female Proposer	0.040	-0.003	-	-	0.039
female Proposer from traditional Village	-0.153**/j-	-	-	-	-0.146**/j-
village Dummies	all***	all***	all***	all	-
R^2-adj.	27.25	26.03	26.36	0	20.63
	Model 6	Model 7	Model 8	Model 9	
constant	0.432***	0.440***	0.919***	0.922***	
offer	0.022***	0.022***	0.018	-	
traditional Village	0.077*	0.243	0.019	-	
female Proposer	-	-	-0.090	-	
female Proposer from traditional Village	-0.107*/j-	-	-	-	
village Dummies	-	-	-	-	
R^2-adj.	20.58	19.19	0	0	

Note: 242 observations.

***/**/*Indicates 1/5/10 percent significance level. j-/ j* indicate joint in/significance of interaction variable and individual interaction terms.

Table A5: *Decision Error and Symmetric Social Preferences in Roth et al. (1991)*

Raw Data / Ungrouped

	N	μ	b	$-\ln(L)$
pie equivalent 10 US$	125	22.419	10.987	859.223
Jerusalem	30	23.723	8.253	207.473
Tokyo	29	29.482	9.727	208.443
Ljubljana	30	19.477	15.385	199.619
Pittsburgh (10 US$)	36	19.358	13.150	239.960

9 Appendix B

Quantal Response Equilibrium for normal-form games (NQRE)

Let $\Gamma = (n, S_i, u_i)$ be a finite n-person game in the normal form. Let $S_i = \{s_{i1}, ..., s_{iJ_i}\}$ denote the strategy set consisting of J_i pure strategies for each player $i \in \{1, ...n\}$. The utility function u_i maps $(S_{1,...,}S_n) \rightarrow \mathbb{R}$. Further, let $\triangle_i$ be the set of probability measures on S_i. Elements of $\triangle_i$ are of the form $p_i : S_i \rightarrow \mathbb{R}$, where $\Sigma_{s_{ij} \in S_i} p_i(s_{ij}) = 1$ and $p_i(s_{ij}) \geq 0$ for all $s_{ij} \in S_i$. Following McKelvey and Palfrey (1995, pp. 9-10), I use the notation $p_{ij} = p_i(s_{ij})$, the 'abusive' notation s_{ij} to denote the strategy $p_i \in \triangle_i$ with $p_{ij} = 1$, and the shorthand notation $p = (p_i, p_{-i})$. Hence, (s_{ij}, p_{-i}) represents the strategy where i adopts the pure strategy s_{ij} and all other players adopt their components of p.

In the utility maximization, players make decision errors because their experienced utility $\widehat{u}_i$ is assumed to be disturbed by additive random shocks:

$$\widehat{u}_i(s_{ij}, p_{-i}) = \overline{u}_i(s_{ij}, p_{-i}) + \varepsilon_{ij},$$

where, according to behavioral assumption $(Q2)$, $\overline{u}_i = E(u_i(s_{ij}, p_{-i}))$ is the expected utility from the 'true' deterministic utility function, and player i's disturbances $\varepsilon_i = (\varepsilon_{i1}, ..., \varepsilon_{iJ_i})$ are outcome specific shocks from a commonly known distribution function $f_i(\varepsilon_i)$ with $E(\varepsilon_i) = 0$. The behavioral assumption $(Q1)$ is that each player selects an action j such that $\widehat{u}_i(s_{ij}, p_{-i}) \geq \widehat{u}_i(s_{ik}, p_{-i})$, $k \in \{1, ..., J_i\}$, i.e. each player chooses action j such that $\widehat{u}_{ij}$ is maximal, given the other players' behavior p_{-i}. For any given $\overline{u}_i$ and $f = (f_1, ..., f_n)$, this decision rule implies a probability distribution σ_i over the strategy set S_i of each player, induced by the probability distribution f over the vector of disturbances ε_i. McKelvey and Palfrey (1995) call this probability distribution $\sigma_i(\sigma_{i1}, ..., \sigma_{iJ_i})$ over the strategy set the quantal response (or statistical reaction) function. Thereby the probability σ_{ij} that player i will select strategy j is given by:

$$\sigma_{ij}(\overline{u}_i) = \int_{R_{ij}(\overline{u}_i)} f(\varepsilon)\, d\varepsilon,$$

where $R_{ij}(\overline{u}_i) = \{\varepsilon_i | \overline{u}_i + \varepsilon_{ij} \geq \overline{u}_i + \varepsilon_{ik} \forall k = 1, ..., J_i\}$ specifies the region of errors that will lead i choose action j and $\varepsilon = (\varepsilon_1, ..., \varepsilon_n)$.

A QRE is defined as a mutual best response in the statistical reaction functions. Formally, a vector $\sigma^* = (\sigma_1^*, ..., \sigma_n^*)$ constitutes a QRE, if for all $i \in \{1, ..., n\}$:

$$\sigma_{ij}^* = \sigma_{ij}(\overline{u}_i(\sigma^*)).$$

A particular QRE is the logit equilibrium. It assumes *logit response functions* $\widetilde{\sigma}_i$ for all players, in which the probability σ_{ij}, that player i will select strategy j, stems from a logistic distribution:

$$\widetilde{\sigma}_{ij}\left(\overline{u}_i\right)=\frac{e^{\lambda\overline{u}_{ij}}}{\Sigma_{k=1}^{J_i}e^{\lambda\overline{u}_{ik}}},$$

where $\lambda\geq 0$ is a non-negative constant. The logit response function corresponds to optimal choice behavior in a QRE, if $(Q3)$ the error distribution f_i has an extreme value distribution with cumulative density function $F_i\left(\varepsilon_{ij}\right)=e^{-e^{-\lambda\varepsilon_{ij}-\gamma}}$ and the ε_{ij}'s are independent. Therefore, if each player uses a logit response function $\widetilde{\sigma}_i$ the logit equilibrium $\widetilde{\sigma}_i^*$ also requires for each player i and strategy j:

$$\widetilde{\sigma}_{ij}^*\left(\overline{u}_i\right)=\frac{e^{\lambda\overline{u}_{ij}}}{\Sigma_{k=1}^{J_i}e^{\lambda\overline{u}_{ik}}}.$$

Thus, in the logit equilibrium the odds are determined by an exponential transformation of the utility times the constant λ. The ratio of probabilities $\widetilde{\sigma}_{ih}^*/\widetilde{\sigma}_{ij}^*$ of two different strategies, s_h and s_j, is given by $\exp\left[\lambda\left(Eu\left(s_h,p_{-i}\right)-Eu\left(s_j,p_{-i}\right)\right)\right]$.[27] The constant λ is called the coefficient of certitude and can be interpreted as a measure of the responder's rationality.[28] As $\partial\widetilde{\sigma}_{ij}^*/\partial\overline{u}_{ij}>0$ the logit equilibrium captures the intuitive notion that better responses are more likely to be observed than worse responses. Best responses are never played with certainty as $\widetilde{\sigma}_{ij}^*<1$ if the number of available pure strategies $J_i\geq 2$.

[27] $\widetilde{\sigma}_{ih}^*\left(\overline{u}_i\right)/\widetilde{\sigma}_{ij}^*\left(\overline{u}_i\right)=\frac{e^{\lambda\overline{u}_{ij}}}{\Sigma_{k=1}^{J_i}e^{\lambda\overline{u}_{ik}}}\frac{\Sigma_{k=1}^{J_i}e^{\lambda\overline{u}_{ik}}}{e^{\lambda\overline{u}_{ih}}}=e^{\lambda\left(\overline{u}_{ih}-\overline{u}_{ij}\right)}$. Behavioral assumption $(Q2)$, which requires $\overline{u}_i=E\left(u_i\left(s_{ij},p_{-i}\right)\right)$, completes the proof. ■

[28] Equivalently a decision error $\mu=1/\lambda$ for $0<\lambda<\infty$ can be used to measure the amount of noise.