

‘Economic Man’ in Cross-cultural Perspective: Behavioral Experiments in 15 Small-scale Societies

Joseph Henrich, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, Richard McElreath, Michael Alvard, Abigail Barr, Jean Ensminger, Kim Hill, Francisco Gil-White, Michael Gurven, Frank Marlowe, John Q. Patton, Natalie Smith, and David Tracer¹

This paper is a contribution to the research of both the Economics Program and the Program on Cultural and Social Dynamics at the Santa Fe Institute. It is under review at Behavioral and Brain Sciences.

Keywords: Self-interest, altruism, cooperation, Ultimatum Game, Public Goods Game, cross-cultural research, experimental economics, game theory

¹ Affiliations: **Henrich**, Wissenschaftskolleg zu Berlin, Wallotstrasse 19, 14193-Berlin, j.henrich@wiko-berlin.de; **Boyd**, Dept. of Anthropology, University of California, Los Angeles CA 90095 rboyd@anthro.ucla.edu; **Bowles**, Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, New Mexico 87501, bowles@santafe.edu; **Camerer**, Div HSS 228-77, Caltech, Pasadena CA 91125 USA, camerer@hss.caltech.edu; **Fehr**, Wissenschaftskolleg zu Berlin, Wallotstrasse 19, 14193-Berlin, wiko1ef@wiko-berlin.de; **Gintis**, Dept. of Economics, University of Massachusetts, Amherst, MA 01060, hgintis@mediaone.net; **McElreath**, University of California Davis, Dept. of Anthropology, One Shields Ave, Davis, CA, rlm@ucla.edu; **Alvard**, Dept. of Anthropology, 4352 TAMU, Texas A&M University, College Station, Texas 77843-4352, Alvard@tamu.edu; **Barr**, Centre for the Study of African Economies, University of Oxford, St. Cross Building, Manor Road, Oxford OX1 3UL, abigail.barr@economics.ox.ac.uk; **Hill**, Dept. of Anthropology, University of New Mexico, kimhill@unm.edu; **Gurven**, Dept. of Anthropology, University of California-Santa Barbara, Santa Barbara, CA 93106, gurven@anth.ucsb.edu; **Patton**, Dept. of Anthropology, Washington State University, Pullman, WA 99164-4910, pattonj@mail.wsu.edu; **Tracer**, Dept. of Anthropology & Health & Behavioral Sciences, University of Colorado at Denver, Denver, CO 80217, dtracer@carbon.cudenver.edu; **Ensminger**, Div HSS 228-77, Caltech, 1200 East California Blvd., Pasadena, CA 91125-7700, jensming@hss.caltech.edu; **Gil-White**, Dept. of Psychology, University of Pennsylvania, 3815 Walnut Street, Philadelphia PA 19104-6196, ffgil@psych.upenn.edu; **Marlowe**, Dept. of Anthropology, Peabody Museum, 11 Divinity Avenue, Harvard University, Cambridge, Massachusetts 02138, fmarlowe@fas.harvard.edu; **Smith**, Dept. of Anthropology, University of California Los Angeles, Los Angeles, CA 90024, nathenrich@hotmail.com

Abstract

Experimental behavioral scientists have found consistent deviations from the predictions of the canonical model of self-interest in over a hundred experiments from around the world. Prior research cannot determine whether this uniformity results from universal patterns of behavior, or from the limited cultural variation among the university students subject pools used in virtually all prior experimental work. To address the above questions, we undertook a cross-cultural study of behavior in Ultimatum, Public Goods, and Dictator Games in fifteen small-scale societies exhibiting a wide variety of economic and cultural conditions.

We found, first, that the canonical selfishness-based model fails in all of the societies studied. Second, there is more behavioral variability than had been found in previous research. Third, group-level differences in economic organization and the structure of social interactions explain a substantial portion of the behavioral variation across societies: the higher the degree of market integration and the higher the payoffs to cooperation in everyday life, the greater the level of prosociality expressed in experimental games. Fourth, individual-level economic and demographic variables do not explain game behavior, either within or across groups. Fifth, in many cases experimental play appears to reflect the common interactional patterns of everyday life.

Short Abstract

Experiments have revealed consistent deviations from self-interest. Existing research cannot determine whether this results from universal human motives or is a property of university student used as subjects. We undertook a cross-cultural experimental study in fifteen small-scale societies. We found that (1) the self-interest model fails in all of the societies studied, (2) there is more variability than had previously been observed, (3) group-level differences explain a substantial portion of the behavioral variation across societies, (4) individual-level economic and demographic variables do not explain behavior, and (5) experimental play often reflects the interactional patterns of everyday life.

In the 13 years since “Selfishness examined...” appeared in these pages, additional experimental evidence has strongly confirmed the doubts expressed by Caporael and her collaborators (1989) concerning the adequacy of self-interest as a behavioral foundation for the social sciences. Experimental economists and others have uncovered large and consistent deviations from the predictions of the textbook representation of *Homo economicus* (Roth 1995, Fehr & Gächter 2000, Camerer 2002). Literally hundreds of experiments in dozens of countries using a variety of game structures and experimental protocols suggest that in addition to their own material payoffs, subjects care about fairness and reciprocity, are willing to change the distribution of material outcomes among others at a personal cost to themselves, and reward those who act in a prosocial manner while punishing those who do not, even when these actions are costly. One of these experiments, the Ultimatum Game, has been implemented in over a hundred studies in two dozen countries with strikingly uniform results. Initial skepticism about the experimental evidence has waned as subsequent experiments with high stakes and with ample opportunity for learning failed to modify the initial conclusions.

These experiments create an empirical challenge to what we call the *selfishness axiom*—the assumption that individuals seek to maximize their own material gains in these interactions and expect others to do the same.² In response there have appeared a

² We extend this formulation of the axiom to cover cases in which individuals maximize the *expected* utility of their material gains to address the question of risk aversion, but use this simpler formulation otherwise.

number of reformulations of individual utility functions and other behavioral foundations consistent with the evidence from across a variety of experimental settings (Fehr & Schmidt 1999, Falk & Fishbacher 1999, Charness & Rabin 1999) as well as attempts to explain the long term evolutionary success of non-selfish behaviors (Simon 1990, Caporael et al 1989, Henrich & Boyd 2001, Boyd et al. 2001, Sober & Wilson 1994, Smith et al. 2001).

Nevertheless, fundamental empirical questions remain unanswered. Do such consistent violations of the canonical model provide evidence of universal patterns of behavior? Or, do individuals' economic and social environments shape their behavior? If the latter, which economic and social conditions are involved? Is reciprocal behavior better explained statistically by individuals' attributes such as their sex, age, and relative wealth, or by the attributes of the group to which the individuals belong? Are there cultures that approximate the canonical account of purely self-regarding behavior?

Existing research cannot answer such questions because virtually all subjects have been university students, and while there are cultural differences among student populations throughout the world, these differences are small compared to the full range of human social and cultural environments. Accordingly, we undertook a large cross-cultural study of behavior in Ultimatum, Public Goods, and Dictator Games. Twelve experienced field researchers (11 anthropologists and one economist), working in twelve countries on four continents and New Guinea, recruited subjects from fifteen small-scale societies exhibiting a wide variety of economic and cultural conditions. Our sample consists of three foraging societies, six that practice slash-and-burn horticulture, four nomadic herding groups and two sedentary, small-scale agricultural societies. Our games

were played anonymously, and for real stakes (the local equivalent of one or more day's wages).

The results of this project, described in detail below, can be summarized in five points: first, there is no society in which experimental behavior is consistent with the canonical model; second there is much more variation between groups than has been previously reported; third, differences between societies in market integration and the importance of cooperation explain a substantial portion of the behavioral variation between groups; fourth, individual-level economic and demographic variables do not explain behavior within or across groups; fifth, experimental play often mirrors patterns of interaction found everyday life. Below we first describe the experimental methods used and give brief descriptions of the societies studied. We then present and interpret our results.

The Cross-cultural Behavioral Experiments Project

Early cross-cultural economics experiments (Roth et. al. 1991 and Cameron 1999) showed little variation among societies: whether in Pittsburgh, Ljubliana, Yogyakarta, or Tokyo, university students played these games in much the same way. However, in 1996 an anomalous experiment finding broke the consensus: the Machiguenga, slash-and-burn horticulturalists living in the southeastern Peruvian Amazon behaved much less prosocially than student populations around the world (Henrich 2000). What then appeared as “the Machiguenga outlier” sparked curiosity among a group of behavioral scientists: was this simply an odd result, perhaps due to the unusual circumstances of the experiment, or had Henrich tapped real behavioral differences, perhaps reflecting the distinct economic circumstances or cultural environment of this Amazonian society? In

November of 1997, Gintis and Boyd, directors of the MacArthur Foundation Research Network on the Nature and Origin of Preferences, invited 12 experienced field workers along with several behavioral economists and experimentalists to a three-day meeting. During this meeting we redesigned the experiments—typically conducted in computer labs at universities—for field implementation often in quite remote areas and among non-literate subjects, and developed common protocols and games stakes for across the many field sites. Two years later, after all the ethnographers had returned from the field, we reconvened to present, compare, and discuss our results. Here we summarize our findings thus far. Planning for a second round of experiments is currently underway.

The Experiments

The field researchers performed three different kinds of economics experiments: Ultimatum Bargaining Games (UG), Public Goods Games (PGG), and Dictator Games (DG). Every field worker did the UG, several administered some form of PGG and three did the DG. Below, we briefly describe these three games, although interested readers should see Kagel & Roth (1995) and Davis & Holt (1993) for details.

The Ultimatum Game

The UG is a simple bargaining game that has been extensively studied by experimental economists. In this game subjects are paired, and the first player, often called the “proposer,” is provisionally allotted a divisible “pie”(usually money). The proposer then offers a portion of the pie to a second person, often called the “responder.” The responder, knowing both the offer and the total amount of the pie, then has the opportunity to either accept or reject the proposer’s offer. If the responder accepts, he or she receives the amount offered and the proposer receives the remainder (the pie minus

the offer). If the responder rejects the offer, then neither player receives anything. In either case, the game ends; the two subjects receive their winnings and depart. Players typically receive payments in cash and remain anonymous to other players, but not to the experimenters (although experimental economists have manipulated both of these variables). In the experiments described here, players were anonymous, and the games involved substantial sums of the appropriate currency. For this game, the canonical assumptions (i.e., all participants maximize their income and this is known by all of them) predict that responders, faced with a choice between zero and a positive payoff should accept any positive offer. Knowing this, proposers should offer the smallest non-zero amount possible. In every experiment yet conducted the vast majority of subjects have violated this prediction.

The Dictator Game

The Dictator Game is played exactly like the standard Ultimatum Game, except that the responder is not given an opportunity to accept or reject the offer. The proposer merely dictates the division. In the Dictator Game positive offers cannot result from a fear of rejection. Thus, when used in conjunction with the Ultimatum Game, this experimental tool allows researchers to determine whether proposers make positive offers out of a 'sense of fairness' or from a 'fear of rejection'.

Public Goods Games

Public goods experiments are designed to investigate how people behave when individual and group interests conflict. We used two variants: the 'Voluntary Contributions' (VC) format and the 'Common-Pool Resources' (CPR) format, the only difference being that in the former subjects may contribute to the common good and in

the latter may refrain from withdrawing from the common resource for private gain. In the VC version, players receive some initial monetary endowment. They then have the opportunity to anonymously contribute any portion of their endowment (from zero to the full endowment) to the group fund. Whatever money is in the group fund after all players have had an opportunity to contribute is increased by 50% (or sometimes doubled), and then distributed equally among all players regardless of their contribution. The payoff structure of the CPR version is identical, except that instead of receiving an endowment, players can make limited withdrawals from the group fund. Whatever remains in the fund (the common pool) after everyone has withdrawn is increased by 50% or doubled, and distributed equally among all group members. The game is not repeated. Selfish subjects may calculate that independently of the actions taken by the other players, contributing as little or withdrawing as much as possible maximizes their monetary payoffs. Free riding is thus the dominant strategy for the selfish subject. Thus, rational selfish players should contribute zero to the group fund (or withdraw their limit in the CPR format).

Ethnographic Description

Figure 1 shows the location of each field site, and Table 1 provides some comparative ethnographic information about the societies discussed here. In selecting these locations, we included societies both sufficiently similar to the Machiguenga to offer the possibility of replicating the original Machiguenga results, and sufficiently different from one another to provide enough economic and cultural diversity to allow an exploration of the extent to which behaviors covary with local differences in the structures of social interaction, forms of livelihood, and other aspects of daily life.

[Figure 1 here]

In Table 1, the column ‘Economic Base’ provides a general classification of the production system in each society. *Horticulturalists* rely primarily on slash-and-burn agriculture, which involves clearing, burning and planting small gardens every few years. All the horticulturalists included in this study also rely on a combination of hunting, fishing and gathering. We have classified the Aché economic base as *Horticulture–Foraging* because they were full-time foragers until about 28 years ago, and still periodically go on multi-week foraging treks, but have spent much of the last few decades as manioc-based horticulturalists. The Au and Gnau of Papua New Guinea are classified as *Foraging/Horticulture* because, despite planting slash and burn gardens, they rely heavily on harvesting wild sago palms for calories, and game for protein. Unlike foragers and horticulturalists, *Pastoralists* rely primarily on herding livestock, often cattle. *Agro-pastoralists* rely on a combination of small-scale sedentary agriculture and herding. We labeled the Orma, Mongols and Kazakhs as pastoralists because most people in these groups rely entirely on herding, although some members of all three groups do some agriculture. The Sangu are labeled *Agro-pastoralists* because many Sangu rely heavily on growing corn, while others rely entirely on animal husbandry. (consequently, in some of our analyses we separate Sangu herders and Sangu farmers).

[Table 1 here]

The column ‘Residence’ classifies societies according to the nature and frequency of their movement. *Nomadic* groups move frequently, spending as little as a few days in a single location, and as long as a few months. *Semi-nomadic* groups move less frequently, often staying in the same location for a few years. Horticultural groups are often semi-nomadic, moving along after a couple of years in search of more abundant game, fish,

wild foods and fertile soils. *Transhumant* herders move livestock between two or more locations in a fixed pattern over the course of a year, often following the good pasture or responding to seasonal rainfall patterns. *Bilocal* indicates that individuals maintain two residences and spend part of the year at each residence. The Machiguenga, for example, spend the dry season living in villages along major rivers, but pass the wet season in their garden houses, that may be located three or more hours from the village. Classifications of the form *Bilocal–Semi-nomadic* indicate that the Machiguenga, for example, were traditionally semi-nomadic, but have more recently adopted a bilocal residence pattern. Similarly, the Aché are classified as *Sedentary–Nomadic* because of their recent transition from nomadic foraging to sedentary horticulture.

The column ‘Language Family’ provides the current linguistic classification for the language traditionally spoken by these societies, and is useful because linguistic affinity provides a rough measure of the cultural relatedness of two groups. The classification of the Mapuche, Hadza, Tsimane and New Guinean languages require special comment. There is no general agreement about how to classify Mapuche with the other language groups of South America—it is often regarded as a linguistic isolate. Similarly, although it was once thought that Hadza was a Khoisan language, distantly related to the San languages of southern Africa, agreement about this is diminishing. The Tsimane language resembles Mosesten (a Bolivian group similar to the Tsimane), but otherwise these two languages seem unrelated to other South American languages (except perhaps distantly to Panoan). Finally, because of the linguistic diversity of New Guinea, we have included both the language phylum for the Au and the Gnau, Torricelli, and their local language family, Wapei.

The column, 'Complexity' refers to the anthropological classification of societies accord to their political economy (Johnson & Earle 2000). *Family-level* societies consist of economically-independent families that lack any stable governing institutions or organizational decision making structures beyond the family. Societies classified as *Family plus extended ties* are similar to family-level societies, except that such groups also consistently exploit extended kin ties or non-kin alliances, for specific purposes such as warfare. In these circumstances, decision making power is ad hoc, ephemeral, and diffuse, but high status males often dominate the process. *Bands* consist of both related and unrelated families that routinely cooperate in economic endeavors. Decision making relies heavily on group consensus, although the opinions of high status males often carry substantial weight. *Clans* and *Villages* are both corporate groups of the same level of complexity, and both are usually larger than bands. *Clans* are based on kinship, tracked by lineal descent from a common ancestor. Decision-making power is often assigned based on lineage position, but prestige or achieved status may play a role. *Villages* operate on the same scale of social and political organization as clans, but consist of several unrelated extended families. Decision making is usually vested in a small cadre of older, high status men who may compete fiercely for prestige. At a larger scale of organization, *Multi-Clan Corporate* groups are composed of several linked clans, and are governed by a council of older high status men—assignment to such councils is often jointly determined by lineal descent and achieved prestige. Multiclan corporations sometimes act only to organize large groups in times of war or conflict, and may or may not play important economic role. Often larger than multiclans corporations, *Chiefdoms* are ruled by a single individual or family and contain several ranked clans or villages.

Rank of individuals and clans/villages usually depends on real or customary blood relations to the chief. Economic organization and integration in chiefdoms is more intense than in multiclan corporate groups, and chiefs usually require subjects to pay taxes or tribute. Such payments allow for the large-scale construction of irrigation works, monuments, and public buildings, as well as the maintenance of standing armies.

The two remaining columns, market integration and payoffs to cooperation, refer to rankings we constructed on the basis of our own and others' ethnographic investigations; we explain these below.

Experimental Results

Because our comparative data on the Ultimatum Game is much more extensive than for the Public Goods and Dictator Games, we primarily focus on the UG results.

Substantial cross-cultural variability

The variability in Ultimatum Game behavior across the groups in our study exceeded that in the entire empirical literature. Prior work comparing Ultimatum Game behavior among university students from Pittsburgh, Ljubljana (Slovenia), Jerusalem, Tokyo (Roth et. al. 1991; Roth 1995; Slonim & Roth 1990) and Yogyakarta (Java, Indonesia; Cameron 1999) revealed little variation between groups. In contrast, figure 2 summarizes our UG results from 15 different societies. While mean UG offers in experiments with student subjects are typically between 43% and 48%, the mean offers from proposers in our sample span a range from 26% to 58%—both below and above the typical behavior (Table 2 presents additional details) While modal UG offers are consistently 50% among university students, our sample modes vary from 15% to 50%.

[Figure 2 and Table 2 here]

The behavior of responders in the Ultimatum Game (Figure 3) is also much more variable than previously observed. In some groups, rejections are extremely rare, even in the presence of low offers, while in others, rejection rates are substantial, including frequent rejections of ‘hyper-fair’ offers (i.e. offers above 50%). Among the Kazakh, Quichua, Aché and Tsimane, we observe zero rejections after 10, 14, 51 and 70 proposer offers, respectively. And, while the offers to the Aché were mostly equitable, 47% of offers to Tsimane and 57% of the offers to Quichua were at or below 30%—yet all were accepted. Similarly, Machiguenga responders rejected only one offer, despite the fact that over 75% of their offers were below 30% of the pie. At the other end of the rejection scale, Hadza responders rejected 24% of all proposer offers and 43% (9/21) of offers 20% and below. Unlike the Hadza and other groups who preferentially rejected low offers, the Au and Gnau of Papua New Guinea rejected both unfair *and* hyper-fair offers with nearly equal frequency, a seemingly odd finding which will presently provide considerable insight into the relationship between experimental behavior and daily life. University student responders fall towards the upper end of the rejection scale (with more rejection than average), but still reject less than some groups like the Au, Gnau, Sangu farmers, and Hadza, all of whom rejected positive offers with greater frequency than (for example) the Pittsburgh subjects in Roth et al. (1991).

[Figure 3 here]

As in the UG, Voluntary contributions (VC) and common-pool resource (CPR) games, which we will collectively call public goods games (PGG), also show much greater variation than previously found in public goods games run in industrialized

societies, and all these results conflict with the predictions of self-regarding models under typical assumptions. Typical distributions of PGG contributions from university students have a ‘U-shape’ with the mode at full defection (those who contribute zero) and a secondary mode at full cooperation (those who contribute everything to the group). The mean contribution is usually between 40% and 60%. Table 3 shows that our cross-cultural data provides some interesting contrasts with this pattern. The Machiguenga, for example, have a mode at full defection, but lack any fully cooperative contributions—which yields a mean contribution of 22%. Both the Aché and Tsimane experiments yielded means between 40-60%, like folks from industrialized societies, but show unimodal distributions with peaks at 50% and 66.7%, respectively. Their distributions resemble *inverted* American distributions with few or no contributions at full free-riding and full cooperation. Like the Aché and Tsimane, the Orma and Huinca have modes near the center of the distribution, at 40% and 50% respectively, but they also show secondary peaks at full cooperation (100%)—and no contributions at full defection. Interestingly, the Orma and Huinca distributions resemble the 1st round of a finite, repeated PGG done with university students (similar to Fehr & Gächter 2000, for example; see Henrich & Smith, this volume).

[Table 3 about here]

Violations of the Selfishness Axiom

In one way or another the Selfishness Axiom was violated in every society we studied across all three different experimental games (DG, UG and PGG). Focusing on the UG, either proposer or responder behavior violated the Axiom, or both. Responder behavior was consistent with selfish motives in several groups, but, like university

students, Au, Gnau, Sangu farmers, and Hadza subjects rejected positive offers contrary to the prediction of the selfishness axiom. However, as shown in Figure 3, responders from the Aché, Tsimane, Machiguenga, Quichua, Orma, Sangu herders, and Kazakhs all have rejection rates of less than 5%, roughly consistent with the canonical model. For some groups these low rejection rates are uninformative because all the offers were near 50% (e.g. the Aché and Sangu), so no one in the group received low offers. However, proposers in several groups provided numerous low offers that were virtually never rejected. The self-interest axiom accurately predicts responder behavior for about half of our societies, even though it generally fails to predict the responder behavior of university student.

Proposer behavior was consistent with income maximizing behavior among only two groups, Hadza and Sangu Farmers. On the receiving end, Among university subjects, it is generally thought that offers are fairly consistent with expected income-maximizing strategies *given* the distribution of rejections across offers (Roth et. al. 1991). This was not the case in most of the groups we studied. In four groups (Aché, Tsimane, Kazakhs and Quichua) we could not estimate the income-maximizing offer (IMO) because there were no rejections. Nevertheless, as discussed above, it seems likely that the substantially lower offers would have been accepted. In two groups (Au, and Gnau) the IMO could not be established because responders from these groups did not preferentially accept higher offers, which is perhaps an even more striking violation of the selfishness axiom.

For each of the remaining societies we used the responder data to estimate the aforementioned IMO.³ Estimates of the IMO are substantially higher than observed mean offers for the all the remaining societies save two. The IMO is a useful measure of the frequency with which low offers were rejected. If rejections are few, or if their likelihood of being rejected is not strongly related to the size of the offer, the IMO will be low (as, for example, is the case for the Tsimane). If substantial offers are frequently rejected, the IMO will be high (e.g. Sangu farmers). Figure 4 compares IMO's (calculated from responder data) to actual mean offers (from proposers). Mean offers by the Hadza and the Sangu (farmers) approximated the IMO (but, in both groups responder behavior violates the Selfishness Axiom). However, for the other groups, mean offers were all substantially above the IMO, ranging from Sangu herders whose mean offers were 130% of the IMO to the Achuar whose mean offers were 400% of the IMO. We conclude that the behavior of proposers in our groups generally does not match the prediction based on the selfishness axiom.

[Figure 4 here]

It is possible that high offers are consistent with a more conventional extension of the selfishness axiom, namely risk aversion. It is a common (though not universal) observation that people prefer a certain amount of money to a gamble with the same expected payoff. Economists model this behavior by assuming that people seek to maximize their expected utility, and that utility is a concave function of income (diminishing returns). For example, suppose a subject estimates that an offer of 40% of

³ Due to the paucity of rejections some of these estimates are not very precise.

the pie will be accepted for sure, and that an offer of 10% will be accepted with probability $2/3$. If she was risk averse, she could value the certainty of keeping 60% of the pie more than the $2/3$ chance of keeping 90% (and a $1/3$ chance of getting nothing). In this case the expected monetary gain is the same for the two offers (60% of the pie), but the *expected utility* of the certain outcome is greater. Thus, a highly risk averse subject might make a high offer even if the probability of rejection of a low offer is small.

There are two reasons to doubt that risk aversion explains proposer behavior in our samples. First, the degree of risk aversion necessary to explain the behavior we observed is much higher than is typically seen in gambles for the kinds of stakes used in our experiments. To determine if utility maximization by risk averse proposers could explain our observations, we transformed the game payoffs into utilities using varying levels of risk aversion, and for each group estimated the degree of risk aversion sufficient that the observed mean offer would be utility maximizing.⁴ The Hadza and the Sangu farmers were approximately expected income maximizers, and thus their offers are consistent with expected utility maximization for risk neutral individuals. But for the other groups—Orma, Sangu herders, Machiguenga, Mapuche, and Shona—the implied levels of risk aversion are implausible. Even for the least extreme case, the Shona, the necessary degree of risk aversion necessary to make their behavior consistent with expected utility maximization implies that they would be indifferent between an even chance that an offer

⁴ See appendix 2 of Henrich, et al. (2002). We modeled risk aversion by expressing a subject's utility as one's payoff raised to the exponent r where an individual for whom $r < 1$ is risk averse, $r = 1$ risk neutral, and $r > 1$ risk preferring. We then found the r for which the observed mean offer maximized the expected utility of the proposers, where the expectation is taken over all possible offers and the estimated likelihood of their being rejected.

of 1 out of 10 dollars would be accepted (an expected payoff of \$4.5) and getting 4 cents with certainty.⁵ Clearly, an individual with this level of risk aversion would be unable to function in an uncertain environment. Second, risk aversion was measured directly in the Mapuche and the Sangu by offering subjects a series of risky choices (Henrich and Smith 2002, Henrich & McElreath in press). In both societies, subjects were risk preferring, not risk averse, a fact that casts further doubt on the risk aversion interpretation. We conclude that our offers are not explained by risk aversion in the usual sense intended by economists

It is quite possible that high offers reflected a desire to avoid rejections in some sense not consistent with the canonical model (for example, fear that a rejection would be considered an insult or a desire to avoid conflict in the group) and we shall return to these possibilities below.

Additional evidence against the selfishness axiom comes from our three Dictator Game experiments: the results here are more transparent than for the UG because the proposer is simply giving money away with no possibility of rejection. In each of the three groups in which the Dictator Game was played, offers deviate from the typical

⁵ Because the numbers of rejections are small, some of our estimates of risk aversion are very imprecise. Accordingly one might worry that more reasonable estimates of risk aversion might fit the data nearly as well as the best fit. To test for this possibility, we computed the difference between the best-fit value of r and 0.81, the value estimated by Tversky and Kahneman (1992) for laboratory data on risky decision making. For some data sets the difference was small, and others quite large. Moreover, there is positive but non-significant correlation between deviation of observed behavior from the IMO and this measure of the precision of the estimate of r . Thus, it seems unlikely that risk aversion can be a complete explanation for our observations.

behavior of university students and from the predictions of self-regarding models. Mean offers among the Orma, Hadza, and Tsimane, respectively, were 31, 20 and 32 percent of the pie. These mean Dictator offers are 70, 60 and 86 percent of the corresponding mean UG offers for these groups. And, few or none of the subjects in these societies offered zero, while the modal offer among university students is typically zero.

Finally, the results from all six of our Public Goods Games also conflict with the selfishness axiom, with means ranging from 22% among the Machiguenga to 65% among the Ache. Except for the Machiguenga (and student populations), no group has more than 5% full defectors.

Explaining Group Differences in Behavior

We first attempted to determine whether any attributes of individuals were statistically associated with proposer offers across our sample. One reflection of the diversity of the societies in our study is the paucity of quantifiable individual level variables that are available and meaningful across the populations we studied. Among the measured individual attributes that we thought might statistically explain offers were the proposer's sex, age, level of formal education, and their wealth relative to others in their group.⁶ In pooled regressions using all offers we found that none of these individual measures predicted offers once we allowed for group level differences in offers (by

⁶ Relative wealth was measured by the in-group percentile ranking of each individual, with the measure of individual wealth varying among groups: for the Orma and Mapuche we used the total cash value of livestock, while among the Au, Gnao and Machiguenga we used total cash cropping land. Estimates using relative wealth were restricted to proposers in the seven groups for which we have wealth data.

introducing dummy variables for each of our groups). Since the group dummies account for about 12% of the variance of individual offers, we conclude that group differences are important. However, for the moment we remain agnostic about the role of individual differences. Our pooled regression tested for common effects of these variables across all the groups and hence does not exclude the possibility that the individual differences we have measured may predict behaviors in different ways from group to group. We return to this possibility below.

We speculated that the large between-group differences in offers might reflect differences among groups in the ways that group-members typically interact in the pursuit of their livelihood, in governance of their common affairs, and in other respects. In our efforts to understand why groups might vary so much in their game play we rank ordered our societies in five categories: 1) *Payoffs to Cooperation*—what is the potential benefit to cooperative as opposed to solitary or family based productive activities? 2) *Market Integration*—do people engage frequently in market exchange? 3) *Anonymity*—how important are anonymous roles and transactions? 4) *Privacy*—how well can people keep their activities secret from others? and 5) *Complexity*—how much centralized decision-making occurs above the level of the household.

For *Payoffs to Cooperation* (PC), groups like the Machiguenga and Tsimane ranked the lowest because they are almost entirely economically independent at the family level—no one's economic well-being depends on cooperation with non-relatives. In contrast, the economy of the whale hunters of Lamalera depends on the cooperation of large groups of non-relatives. For *Market Integration* (MI) Hadza were ranked low because their life would change little if markets suddenly disappeared. Others, like the

Orma herders are ranked higher because they frequently buy and sell livestock and work for wages. *Anonymity* (AN) deals with the prevalence of anonymous roles in a society. While many Achuar of the Ecuadorian Amazon never interact with strangers, the Shona of Zimbabwe frequently interact with people they do not know and may never see again. Our measure of *Privacy* (PR) captures the fact that in groups like the Au, Gnaou and Hadza, who live in small villages or bands and eat in public, it's nearly impossible to keep secrets and it's quite difficult to hide anything of value. Among the Hadza, simply having pants substantially increases privacy because they have pockets (which is a reason for their popularity among some Hadza). In contrast, Mapuche farmers live in widely scattered houses and maintain strict rules about approaching another's house without permission, so privacy is substantial.

Before we began the collective analysis we ranked the groups along these dimensions using the following procedure: During a meeting of the research team, we had a lengthy discussion of the underlying attributes that each dimension was designed to capture. Then the field researchers lined up and sorted themselves by repeatedly comparing the characteristics of the group that they studied with their two neighbors in line, switching places if necessary, and repeating the process until no one wanted to move. The subjective nature of the resulting ordinal measures is quite clear.⁷

We assume that these indices are exogenous in the sense that the behavioral patterns that generate by our experimental subjects not also causes of the aspects of groups we have captured in our indices. It is for this reason, for example, that we sought

⁷ This procedure was suggested by Abigail Barr who had used it in her field work.

to measure the *potential* payoffs to cooperation—viewed as a characteristic of the local ecology rather than the amount of cooperation actually practiced which depends on choices of the inhabitants. While plausible, this assumption could be false. Societies adhering to a norm of egalitarian sharing, for example, often sustain the custom of eating in public, a practice that makes the food sharing process transparent, minimizes monitoring costs, and reduces the likelihood of conflicts over divisions. Thus, across a sample of groups, generous proposer offers reflecting a group norm of sharing might vary inversely with the degree of privacy as we have measured it, but the causal relationship would be from the sharing norm to privacy rather than the reverse.

We estimated ordinary least squares regression equations for explaining group mean UG offer, using the rankings described above as well as other group characteristics. Two variables—payoffs to cooperation and market integration—are strong predictors of mean offers. Both their normalized regression coefficients are highly significant and indicate that a standard deviation difference in either variable results in about one half or more standard deviation difference in the group mean offers (Figure 4): $\beta_{PC} = 0.72$ ($p = 0.0004$) and $\beta_{MI} = 0.49$ ($p = 0.007$). Together these two variables account for 66% of the variation among societies in mean UG offers. In other regressions we found that offers are lower in more anonymous societies and those with greater privacy, but the coefficients of these variables are smaller and significance levels of these predictors are marginal. If we use the IMO as a predictor of the UG offers along with PC and MI, its coefficient is small (in magnitude), negative and insignificant, while the coefficients of PC and MI remain large and close to significance, suggesting that the effects of economic

structure and cultural differences captured by PC and MI do not substantially influence offers through the IMO. [Figure 4 here]

Using the ranking and other group level variables to estimate regression equations to explain IMO ($n = 9$) we found only the coefficient of PC was both large and significant. A change of one standard deviation in PC results in about a two-thirds change in the standard deviation of the IMO. Thus, payoffs to cooperation is a strong predictor that low offers will be rejected. MI's coefficient was fairly large, but not well estimated.

Because of the importance in the anthropological literature of the conventional classifications of societies by their political complexity (Johnson & Earle 2000)—from family level societies through chiefdoms— we wanted to know whether this variable had explanatory power⁸. Our measure of complexity is statistically indistinguishable from our measure of market integration ($r = 0.91$) and, like market integration, is highly correlated with settlement size. We are thus unable to statistically separate the causal importance of the two. If we consider market integration and complexity as exogenous, and settlement size as a result of one or both of these, then it appears that the effects of market integration or complexity operates in part through the intervening causal effects of settlement size: groups with larger settlement size made higher offers, and preferentially rejected low offers (resulting in a higher IMO).

⁸ Our complexity rankings were generated by both Henrich (who was not blind to our experimental results) and an outside expert on societal complexity (Allen Johnson) who was blind to our results. Henrich's and Johnson's rankings correlated 0.9, and explain nearly identical amounts of the variation in mean UG offers.

Our analysis of the individual level responder data across all groups reveals some of the same basic patterns observed in the proposer data. A responder's age, sex, and relative wealth does not affect an individual's likelihood of rejecting an offer. What does matter is the proportion of the stake offered, and the responders' ethnolinguistic group.

Explaining Individual Differences within groups

In contrast to the remarkable power of our group level measures in statistically explaining between group differences in experimental behaviors, our individual level variables explain little of the variation between individuals in experimental play. With a few group-specific exceptions, nothing that we measured about the individuals other than their group membership (or village, camp, or other subgroup membership) predicted experimental play. It is possible, of course, that the unexplained within group variance in experimental behaviors reflects subjects' lack of comprehension of the game or errors in experimental play that are unrelated to measures like age, education or wage labor participation. We return to this issue when we discuss concerns about our experimental methods. Here we summarize our findings concerning individual attributes and experimental play.

Sex, wealth, and age do not account for any significant portion of the variance in game play within groups. However, sex was marginally significant among the Tsimane, where males offer 10% more than females. And among the Hadza, women's UG offers strongly increased with camp population size, but camp size was not important to men's offers. Conversely, in the Dictator Game, it was the offers of Hadza men that increased with camp size—although this may be an artifact (Marlowe 2002). As in the UG, Public Goods Game data from five societies also reveal no significant effects of sex,

except among Aché men who contribute a bit more than women. Similarly, wealth, in any form (e.g. cash, cows, land), does not predict game behavior. In several circumstances, multiple measures of wealth (e.g. animal wealth, cash, and land-wealth) were gathered and analyzed, as well as an aggregate measure. In these within group analyses, wealth emerged as significant only once in 12 different data sets (including both UG and PGG datasets). The exception arises from an all-male Public Goods Game among the Orma. Controlling for age, education, income, and residence pattern (sedentary vs. nomadic), wealth was the only significant predictor of contributions in a multivariate linear regression, with a standard deviation difference in wealth predicting well over half a standard deviation difference in contributions—we make sense of this finding below.

Several researchers also gathered and analyzed measures of the number of years of formal schooling subjects had.. Analyzing UG data from the Sangu, Orma, Mapuche, Au and Gnau, we find that the extent of schooling does not account for any significant portion of the variation in offers in either bivariate analyses or multivariate regression that controlled for sex, age and wealth. Among the Tsimane, the extent of formal education emerges as marginally significant in a multivariate regression involving age, village, sex, Spanish-speaking ability, trips to the nearest market town, and wage labor participation. More educated Tsimane offer less in the UG game. However, we find no effect of formal education on PGG play in the Tsimane. Thus while schooling effects may exist in a few cases, they are not particularly strong or consistent across games or societies.

Although our group level measure of market integration has impressive statistical power, individual level measures of market exposure do not explain any significant proportion of the variation within groups. To assess market exposure, some of us gathered data on individuals' participation in wage labor, their reliance on cash cropping, and their competence in the national language. Wage labor participation shows no significant relation to offers in the UG in six groups—the Tsimane, Aché, Gnau, Au, Machiguenga and Mapuche. In these groups, individuals who participate in wage labor make offers that are indistinguishable from those who do not. PGG data from the Orma, Aché, Machiguenga and Tsimane also indicate that wage labor does not affect game play. The only clear exception to the wage labor pattern occurs in the Orma UG data, where individuals who have participated in wage labor make significantly higher offers than those who had not.

In societies based on agriculture, another measure of market integration is the amount of land an individual (or household) devotes to cash cropping, as opposed to subsistence cropping. We have cash cropping data from three societies. Among the Machiguenga, land (in hectares or as a proportion of total land) devoted to cash cropping is highly correlated with UG offers; its normalized partial regression coefficient when age, sex and wage labor are controlled remains substantial, though its significance level is marginal. Neither cash cropping land or the proportion of land devoted to cash cropping is not significantly related to UG offers for the Au and Gnau. However, among the Au (but not the Gnau) multivariate regressions show that land devoted to *subsistence* cropping positively predicts UG offers, controlling for sex, age, cash cropping land and wage labor.

In many places, an individual's degree of competence in the national language may also represent a measure of market integration, or at least of market exposure. Unfortunately, we only have language data from one society, the Tsimane. Comparing the sample of the most fluent Spanish speakers (who are also the most likely to be educated outside the village) against all others, more fluent speakers offered more in the UG than less fluent speakers. However, using multivariate regression to control for village membership, sex, age, visits to San Borja, years of formal education, and participation in wage labor, we find no relationship between Spanish-speaking ability and UG offers. Furthermore, in the Tsimane PGG, competence in the national language also does not predict contributions, using the same controls.

As is the case for all of our individual level data, except for age and sex, these measures capture individual behaviors that may well be endogenous with respect to the beliefs or preferences our experiments measure. Because it is possible that these measures are the consequence rather than the cause of individual behavioral differences, we were also able to use geographical measures of proximity to market opportunities as exogenous instruments for measuring market exposure in three groups: Tsimane, Au and Gnau. However none of these were significant predictors of proposer behavior.

Given that we sought individual level statistical associations for a number of variables in 15 societies and found just a handful of estimates suggesting substantial effects, we conclude that, other than group membership, the individual-level facts we have collected about our subjects do not consistently predict how individuals will behave. This does not mean that within group variation in subjects' behavior cannot be explained; rather it suggests that the explanation may be group-specific and that we may not have

collected the appropriate individual information. It is also possible that variation within groups is explained by individual genetic differences uncorrelated with our regressors (Plomin 1997), even though variation between groups probably results entirely from economic, social or cultural differences.

Local Group Effects

Our analysis suggests that group-effects may be important, and this opens the question of how to define a group. In the above analyses, we used ethnolinguistic markers to define group membership, but non-ethnolinguistic regional groupings, or smaller local groupings (e.g., villages) may be more appropriate. Our data allow several comparisons. Such small-scale tests allow us to control for a number of variables, including climate, language, regional/national economy, local buying power of the game stakes, and local history. In the Bolivian Amazon, the effects of market integration and local-groups were explored by performing the UG and PGG in five different communities at different distances from the market town of San Borja, the only source of commercial goods, medicines, and wage labor opportunities. Like the Machiguenga, the Tsimane live in small communities scattered along a major riverine drainage system. In this situation, physical distance (in travel time along the river) from San Borja acts as an exogenous proxy measure for the extent of market contact of different Tsimane communities. The results indicate that a community's distance from San Borja is unrelated to UG or PGG behavior. Interestingly, the best predictor for UG proposer behavior and PGG contributions is what community one is from, *independent* of the community's distance from San Borja and population size. So, where a Tsimane lives matters, but small differences in both individual-level measures of market integration and community-level

market variables apparently do not. Among the Tsimane, the relevant group for predicting UG and PGG behavior appears to be smaller than the ethnolinguistic group.

As with the Tsimane, we were surprised to find a number of cases in which group membership effects were strong even in the absence of geographical isolation, suggesting that the processes that generate maintain behavioral differences among groups are can maintain differences between frequently interacting, intermixing and even intermarrying groups. In Chile, Mapuche farmers and non-Mapuche Chilean townspeople, locally called Huinca, have lived side-by-side, intermarried and interacted for about 100 years. Yet, the Mapuche and the Huinca behave quite differently in a single-shot PG game. The Mapuche contributed a mean of 33% to the pot, while the Huinca offered an average of 58%. Moreover, in Ecuador the Achuar and Quichua of Conambo, who interact and intermarry frequently, play the UG quite differently—Achuar proposers offered a mean of 43% while Quichua proposers offered only 25%. This difference is especially notable as Quichua and Achuar subjects were randomly paired, so the proposers from the two groups faced the same probability of rejection. As mentioned above, the single biggest predictor of both UG and PGG offers among the Tsimane was village membership. In Tanzania, Hadza from the biggest camp (which was three times larger than the next largest camp) played the UG much more like university students than like Hadza from the four smaller camps, despite the fact that camps are ephemeral social units and camp membership is quite fluid. For the Hadza, camp population size turns out to be the best predictor of UG offers—the larger the camp, the higher the mean UG offer. Finally, although Sangu herders and farmers make similar UG offers, farmers reject offers

significantly more frequently than herders. Yet, Sangu often change from farmer to herder and back again in the course of one lifetime.

Interestingly however, in some of our other research locations group membership displayed no predictive power. In Mongolia, Torguud Mongols and Kazakhs are separated by deep cultural and historical differences, yet they play the UG similarly. In Papua New Guinea the Au and Gnau, who speak mutually unintelligible languages and show differing degrees of market incorporation, played the UG in the same unusual manner (making and frequently rejecting offers of more than half the pie). In Zimbabwe, resettled Shona live in villages that were made up of strangers at their inception two decades ago, while unresettled Shona live in villages comprised of families that have lived side-by-side for generations. Nonetheless, there are only slight differences in UG behavior among resettled and unresettled groups.

In general, the micro-level variation we observed contrasts with the UG results from the U.S. and Europe in which university students, who speak different languages and live thousands of miles apart, behave quite similarly. Of course, it is possible that variation exists within contemporary societies, but this variation is not represented in university populations. However, recent experiments with subjects outside of universities in western societies have thus failed to uncover behavioral patterns in the UG much different from those observed among university students, although Dictator Game behavior appears quite different (Smith 2001, Burks et al. 2001).

Experimental Behavior and Everyday Life

The fact that group level measures of economic and social structure statistically explain much of the between group variance in experimental play suggests that there might be a relationship between behavior in our games and common patterns of interaction in daily life. In a number of cases the parallels are quite striking, and in some cases our subjects readily discerned the similarity, and were able to articulate it. The Orma, for example, immediately recognized that the PGG was similar to the *harambee*, a locally-initiated contribution that Orma households make when a community decides to construct a public good such as a road or school. They dubbed the experiment “the *harambee* game” and gave generously (mean 58% with 25% full contributors). Recall that among the Au and Gnau of Papua New Guinea many proposers offered more than half the pie, and many of these offers were rejected.

The rejection of seemingly generous offers may have a parallel in the culture of status-seeking through gift giving found in Au and Gnau villages, and throughout Melanesia. In these groups, accepting gifts, even unsolicited ones, implies a strong obligation to reciprocate at some future time. Unrepaid debts accumulate, and place the receiver in a subordinate status. Further, the giver may demand repayment at times, or in forms (political alliances), not to the receiver’s liking—but the receiver is still strongly obliged to respond. Consequently, excessively large gifts, especially unsolicited ones, will frequently be refused because of concern about the obligation to reciprocate.

Among the whale hunting Lamalera of Indonesia, 63% of the proposers in the Ultimatum Game divided the pie equally, and most of those who did not offered more than half (the mean offer was 0.57 of the pie). In real life, when a Lamalera whaling crew

returns with a whale or other large catch, a specially designated person meticulously divides the prey into pre-designated parts allocated to the harpooner, crewmembers, and others participating in the hunt, as well as the sail maker, members of the hunters' corporate group, and other community members (who make no direct contribution to the hunt). Because the size of the pie in the Lamalera experiments was the equivalent of ten days wages, making an experimental offer in the UG may have seemed similar to dividing a whale.

Similarly, in Paraguay the Aché regularly share meat. During this sharing, the hunters responsible for the catch commonly forgo their share, while the prey is distributed equally among all other households. There is no consistent relationship between the amount a hunter brings back and the amount his family receives (Kaplan & Hill 1985) And, successful hunters often leave their prey outside the camp to be discovered by others, carefully avoiding any hint of boastfulness. When asked to divide the UG pie, Aché proposers may have perceived themselves as dividing the game they or a male member of their family had acquired, thereby leading 79% of the Aché proposers to offer either half or 40%, and 16% to offer more than 50%, with no rejected offers.

By contrast, the low offers and high rejection rates of the Hadza, another group of small-scale foragers, are not surprising in light of numerous ethnographic descriptions of these people (Woodburn 1968; Marlowe 2002, Blurton-Jones, personal communication). Although the Hadza extensively share meat (and other foods to a lesser degree), they do not do so without complaint; many look for opportunities to avoid sharing. Hunters sometimes wait on the outskirts of camp until nightfall so they can sneak meat into their shelter without being seen. It seems the Hadza share because they fear the social

consequences that would result from not sharing. Cooperation and sharing is enforced by a fear of punishment that comes in the form of informal social sanctions, gossip, and ostracism. Many Hadza proposers attempted to avoid sharing and many of them were punished by rejection. Thus, we find two foraging peoples—the Aché and the Hadza—at opposite ends of the UG spectrum in both proposers' offers and responders' rejections; their contrasting behaviors seem to reflect their differing patterns of everyday life, not any underlying logic of hunter-gatherer life ways.

Similarly, the life ways of our two family level societies are reflected in their game behavior. Both the Machiguenga and Tsimane live in societies with little cooperation, exchange or sharing beyond the family unit. Ethnographically, both show little fear of social sanctions and seem to care little about public opinion. The Machiguenga, for example, did not even have personal names until recently—presumably because there was little reason to refer to people outside of one's kin circle. Consequently, it's not surprising that in anonymous experimental interactions both groups made low UG offers. Given the Tsimane UG offers vary from village to village, it would be interesting to know if these differences reflect village-level differences in real prosocial behavior.

Like many other small-scale agriculturalists, the Mapuche's relations with their neighbors are characterized by mutual suspicion, envy, and fear of being envied. The Mapuche believe that illness, death and bad luck are caused by the malevolent magic of spiteful neighbors and acquaintances, or sometimes merely by the unintentional power of envious others. Material wealth and good fortune result from trickery, taking advantage of others and making deals with spirits—not from hard work, courage or intelligence. Households keep secrets if they can, and many norms are maintained by fear of social

sanctions, not general goodwill. This pattern of social interaction and cultural beliefs is consistent with the Mapuche's post-game interviews in the UG. Unlike UCLA students, Mapuche proposers rarely claimed that their offers were influenced by a sense of fairness. Instead, most proposers based their offers on a fear of rejection. Even proposers who made hyper-fair offers claimed that they feared rare spiteful responders, who would be willing to reject even 50/50 offers.

Discussions of experimental behavior and everyday life commonly address the real world predictive power of experimental play (Loewenstein 1999). Our concern here has been more modest: to determine if there might be analogous patterns of behavior in the experiments and in the daily life of our subjects. In many societies it appears that there are and that our subjects were aware of the parallels in some cases. But this modest observation begs the causal question: why did our subjects behave as they did?

Discussion: Research Methods

It is possible that the diversity of behaviors we observe is an artifact of our experimental methods in these unusual settings. The problems we faced in this respect are different in degree, not in kind, from those confronting any attempt to make inferences about behavioral patterns from experimental data. We were especially mindful of the fact that individual differences in experimental play may arise from a combination of dispositional differences and differences in the framing effect of the experimental situation itself. These framing effects may have been quite strong in our case because of the oddity of the experimental situation to most of our subjects, who have had little experience with abstract games. Moreover, for many of our subjects it is unusual to interact with anyone from outside their own ethnolinguistic community, as the

experimenters were. Although the considerations raised below highlight some of the difficulties of cross-cultural experimental work, we think the experienced fieldworkers who administered the experiments anticipated and addressed these difficulties, for the most part.

First, the administration of the experiments in novel settings may have given rise to misunderstandings, often rooted in different implicit assumptions. For example, in a pre-game pilot study, some Mongolian subjects believed that by accepting UG offers they would be taking money away from the experimenter, while other subjects, even after being clearly told otherwise, did not believe that they would actually be paid real money. The Mongolian results reported here are from a second round of experiments in which these confusions were eliminated by painstaking, repeated instruction and testing. In most cases experimenters tested subjects for game comprehension before the experiments were implemented, and excluded those who had difficulty grasping the game. In several cases, experimenters used post-game interviews to probe for possible misunderstandings and faulty assumptions. Among the Mapuche, players were ranked according to how well they understood the strategic nature of the game, and how well they were able to do the mathematical calculations involved. After excluding those with inadequate understanding and computational competence, the behavior of the remaining players was independent of the rankings. Similarly, among the Hadza, each player was scored according to the number of practice examples it took for them to learn the game (i.e. give correct answers to hypothetical test examples). Among Hadza men this measure is unrelated to both UG proposer and responder behavior, but for women comprehension is positively and significantly correlated with offer size. We do not know if the covariation of

comprehension and experimental behavior among Hadza women represents the effect of comprehension per se, and hence, represents a problem of experimental design or implementation, or results from the association of comprehension with other correlates of game play for women, such as camp size (a strong predictor of Hadza women's offers).

Another methodological problem in interpreting the cross-cultural results comes from possible experimenter bias. In several cases, the relationship between the experimenter and the participants is much closer, more personal, and longer lasting than in typical university-based experiments. Consequently, it is possible that ethnographers may bias the results of these experiments in different ways than experimenters usually affect the results. Henrich (2000) attempted to control for some of this effect by replicating the Machiguenga UG protocol with UCLA anthropology graduate students. In this control, Henrich and his subjects were all known to one another, had interacted in the past, and would interact again in the future. His results were quite similar to typical UG results in high stakes games among university students, and substantially different from the Machiguenga. This is certainly not a complete control for experimenter bias, but it does control for some elements of the bias. To test for experimenter bias across our samples, we examined the relationship between the time each experimenter had spent in the field prior to administering the games and the mean UG of each group, but found no consistent pattern in the data. Nonetheless, we cannot entirely exclude the possibility that some of the observed between-group differences are the result of differences among the experimenters and the manner in which the experiments were implemented. Our next round of experiments further addresses this concern.

Third, the fact that most of, but not all, our experiments were played for money is likely to have affected experimental behavior. In most societies, money is a powerful framing device: the fact that substantial sums of money are changing hands is a strong cue about the nature of the interaction. We see no reason to think that our subjects were any less eager to pick up cues about appropriate behavior in these experimental situations than university students. In Lamalera, for example, packs of cigarettes were used instead of money to avoid the appearance of gambling—cigarettes are highly valued and can be exchanged for money or favors. We do not know if the many hyper-fair offers made by these whale hunters would have been observed had the pie been denominated in money, or how experimental play might have been affected had the pie been denominated in whale meat. Ethnographic evidence suggests that distinct sharing rules pertain to different goods—meat and honey are meticulously shared among the Aché for example, but goods purchased with money and manioc are not. Experimental play with university students and other data suggest that the means by which a valued resource is acquired influences how it is divided, perhaps because different means of acquisition cue different sharing rules. Goods acquired by chance may be governed by sharing rules that do not apply to goods acquired by labor; and, it seems likely that the experimental pie would be seen as a good acquired by chance. Our subsequent work will explicitly examine the effect of different mediums of exchange on game play.

Fourth, some ethnographers had to modify the standardized game procedures. Three researchers instructed their subjects in large groups on how to play the games, rather than in the one-on-one scenarios employed by the other ethnographers (note, this variant makes no difference for university students; Henrich 2000; Henrich & Smith

2002). In another case, to facilitate the game explanation, Hill used explicit analogies to real-life social interactions to clarify the games.

And Fifth, in an effort to collect rejection data, Alvard and Gil-White made sham offers to responders, instead of presenting the actual proposers' offers. It's unclear how these methodological differences might have influenced the overall results, although among U.S. university students, behavior in both the UG and PGG is not very sensitive to such methodological modifications.

Some have suggested that the common violations of the canonical model in these one shot games arose because the subjects simply had no experience with one shot interactions in their own lives, and thus treated these games *as if* they were repeated. Had the subjects interpreted these experiments in this way, they might have imagined being in the role of responder in some subsequent round, possibly paired with the same partner, and made generous offers (or rejected low offers) to affect the subsequent behavior of this imagined future partner. However, we do not find this interpretation compelling for several reasons. First, extensive post-game interviews by several of our researchers indicate that our subjects did comprehend the one-shot aspect of the games. Second, in some experimental comparisons between one-shot and repeated games, most university students demonstrate clear strategic adjustments as they move from one-shot to repeated contexts (e.g., in gift exchange games, Gächter & Falk 2001), indicating that they can perceive a difference—although this does not occur in the UG (Roth et. al. 1991). Nevertheless, this shows that subjects can, in general, recognize the difference between one-shot and repeated games. Third, when opportunities for reputation building are incorporated into a series of one-shot UG plays, university students make predictable

strategic adjustments compared to a series of one-shot games without reputation building (Fehr & Gächter 2001), which again indicates that they can perceive the difference. Finally, it is important to remember that half of our societies generated mean UG offers between the predictions of the canonical model (near zero) and university students. So, if people make generous offers in one-shot games because they believe (in some sense) that they are playing a repeated game, then university students must understand the one-shot nature of the game *less* than the uneducated people in our small-scale societies. Such a claim would be particularly odd, given that university students participate in real one-shot interactions much more than anyone in our sample. For these reasons we believe our results are neither experimental artifacts nor were they caused by our subjects' inability to distinguish between one shot and repeated interactions.

Discussion: Interpretation

Our data suggest that these between-group behavioral differences, which all violate the selfishness axiom, are the product of the patterns of social and economic interaction that frame the everyday lives of our subjects. There are at least four ways that patterns of social interaction could have these effects (Bowles 1998).

Task performance effects. Economic and social institutions structure the tasks people must perform to make a living, and to remain in 'good standing' in the relevant community. There is ample evidence from experiments, industrial sociology, and ethnography, that commonly performed tasks affect values, and that these values are generalized far beyond the immediate domains of task performance. In experimental work, Sherif (1937) and others have shown that the performance of cooperative tasks (in which success depends on the efforts of many and the rewards are shared) induces

positive sentiments toward those with whom one cooperates, while competitive tasks produce the opposite effect. And, from sociology and ethnography, the degree of autonomy one exercises in making a living, for example, is strongly associated with child rearing values in industrial societies (Kohn 1990) and simple societies (Barry, et al 1959).

Framing and situational construal. Economic and social institutions are situations in the social psychological sense and thus have framing and other situation construal effects (Ross & Nisbett 1991). Economists typically represent a choice situation by a set of feasible actions, beliefs concerning the consequences of actions, and an evaluation of the consequences according to exogenous preferences. But the institutions that define feasible actions may also alter beliefs about consequences of actions and the evaluation of these consequences. For example, a market-oriented society may develop distinct cognitive capacities and habits. The fact that almost everything has a price in market-oriented societies provides a cognitive simplification not available to people in societies where money plays a lesser role: namely allowing the aggregation of disparate objects using a monetary standard as in “\$50 of groceries”. To take another example, extensive market interactions may accustom individuals to the idea that interactions with strangers may be mutually beneficial. By contrast, those who do not customarily deal with strangers in mutually advantageous ways may be more likely to treat anonymous interactions as hostile, threatening, or occasions for opportunistic pursuit of self interest. Recent experiments in industrial societies have shown that contextual cues can significantly change the contributions of undergraduates in social dilemmas. For example, Pillutla and Chen (1999) used two versions of a PGG—one construed as a joint investment and the other as a contribution to a social event. Players contributed

significantly more to the social event than to the investment despite the fact that the two versions had the same payoff structure. Similarly, Hayashi et. al. (1999) show that simple framing differences strongly affect rates of cooperation in an otherwise identical two-person prisoner's dilemma, and that these effects depend on whether one is from Japan or the U.S.

Relationship specific investments. The structure of social interactions affects the benefits and costs of reputation building and other relationship-specific investments and thereby alters the evolution of common norms and the degree of social ties. Societies differ markedly in the frequency of interaction with known individuals and the degree to which interactions are governed by complete contracts as opposed to informal guarantees related to trust and reputation. We know from experiments, for example, that trust and interpersonal commitment often arise where contracts are incomplete, but not under complete contracting (Kollock 1994; Brown, Falk and Fehr 2001); these patterns appear to be replicated in actual exchange situations such as the international diamond market (Bernstein 1992) and the market for raw rubber in Malaysia (Siamwala 1978). If trust and commitment are important parts of one's livelihood, these sentiments may be generalized to other areas of life or evoked in situations which appear similar to everyday life.

Effects on the process of cultural transmission. The structure of social interactions affects the process of cultural learning, as it affects who meets particular cultural models, under which conditions, and with what information about the available behavioral alternatives, their prevalence in the relevant social group, and the degree of success or other experiences of those following differing behavioral rules. For example, in societies in which schooling plays a significant role in child rearing teachers are "high prestige"

cultural models very often representing the behavioral patterns of a socially dominant group, while in societies in which schooling plays a lesser role the cultural models may be more locally representative and dispersed.

Our interpretation of these cases reflects a unified underlying causal model in which preferences and beliefs are endogenous. According to our view, behaviors in a given situation are the result of individuals' *beliefs* about the relationship between actions and consequences and the *preferences* with which they evaluate these consequences. The structure of everyday social interactions affects both beliefs and preferences. The reason is that who we meet when we do particular tasks with particular payoffs influences both the kinds of information we deploy when we update our beliefs and the experiences that lead us to reaffirm or revise our preferences.⁹ The updating of beliefs and preferences may respond to the relative payoffs of those holding distinct beliefs and preferences—the successful may be copied. Or, it may be sensitive to the frequency with which one imitates individuals holding distinct beliefs and preferences—learning may be conformist. In combination, such forms of social learning, as well as individual learning, will produce groups with different combinations of beliefs and preferences (which can occur even in the absence of structured social interaction).

We are convinced that local economic and social structures are reflected in the experimental behaviors we observed, and we think it is reasonable that the connection between local conditions and behaviors can be illuminated by the learning model sketched above. However, we are unclear about some important details of how local

⁹ For a more extended discussion see Bowles (1998), Boyd and Richerson, (1985), and the works cited therein.

situations influence behaviors. Two plausible interpretations come to mind. Perhaps different social and physical environments foster the development of differing generalized behavioral dispositions that are applicable across many domains, as might be the case using the above reasoning concerning task performance or investment in reputation building. For example, Lamalerans may be generally more ‘altruistic’ or ‘fair-minded’ than Machiguenga or Quichua. In our experimental situations, such dispositions could account for the statistical relationships between group economic and social characteristics and experimental outcomes. In contrast, our abstract game structures may cue one or more highly context specific behavioral rules, as is suggested by the situational framing examples given above concerning the use of money. According to this interpretation our subjects were first identifying the kind of situation they were in, seeking analogues in their daily life, and then acting in an appropriate manner. In this case, individual differences result from the differing ways that individuals frame a given situation, not from generalized dispositional differences. The diverse societies in our sample clearly differ markedly in their everyday analogues to the experimental situation, and this would explain both the magnitude of group differences and the statistical association between group level economic and social structure and experimental behavior.

These two approaches are difficult to distinguish empirically, and our dataset does not help us judge their relative importance. But in at least one set of our experiments, the two interpretations support quite different sets of predictions. The context specific approach predicts that behavior when playing in different games (e.g. UG and PGG) will be similar if the game seems similar to the subjects—such that the different games cue

the same behavioral rules. By contrast, the dispositional approach predicts similarity of play in games in which a particular disposition would influence play. If situational cues explain experimental play, we might not observe any correlation between subjects' offers unless the two games evoked the same situational cues in the subjects. It is generally difficult to derive any testable hypotheses from this reasoning in part because the cueing process is obscure.

However, one of our cases allows an illuminating distinction between the two. Recall that the Orma made a connection between the PGG and their customary practice, the *harambee*. The Orma believe that wealthy households should make larger contributions to the *harambee* than poor households. The Orma did not perceive a similar connection between the *harambee* and the UG. Multivariate regressions involving wealth, age, education and income indicate that wealth is the only significant predictor of PGG contributions. The more wealth a person has the more they contribute to the common pool, just like in the real *harambee*. Wealth, however, is not a significant predictor of UG offers in either multivariate or bivariate analyses. The importance of wealth for PG games, but not for UGs, is consistent with predictions from the context specific approach.

The many other cases in which one or more of our experimental situations appeared similar to common social interactions, do not allow us to distinguish between the dispositional and situational interpretations.

Conclusion

We summarize our results as follows. First, the selfishness axiom is not supported in *any* society studied, and the canonical model fails in variety of new ways. Second,

there is considerably more behavioral variability across groups than had been found in previous research. Third, group-level differences in economic organization and the degree of market integration explain a substantial portion of the behavioral variation across societies: the higher the degree of market integration and the higher the payoffs to cooperation, the greater the level of prosociality found in experimental games. Fourth, individual-level economic and demographic variables do not explain behavior either within or across groups. Fifth, behavior in the experiments is generally consistent with economic patterns of everyday life in these societies.

We believe that the degree of variability observed in our cross-cultural sample of societies, and the persistent failure of the selfishness axiom, bears directly on related research emerging in economics (Fehr & Gächter 2000), economic sociology (Kollock 1994), and political science (Ostrom 1998, 2000). In economics, for example, the building blocks of new theories posit preferences such as a sense of fairness, a devotion to reciprocity, an aversion to inequality, a concern for relative payoffs, and a taste for punishment (e.g. Fehr & Schmidt 1999; Charness & Rabin 1999; Bolton & Ockenfels 1999). However, our results suggest that the student populations examined by most experimental social scientists may represent a very limited sample from a quite diverse population of human societies.

It is tempting to react to the widespread experimental evidence of non-selfish behaviors by replacing the *selfishness axiom* with some equally simple and universal assumption about human behavior. If *Homo economicus* has failed the experimental test, maybe *Homo altruisticus*, *Homo reciprocans*, or some other simplified version of a panhuman nature will do better. The diversity of behaviors we have observed leads us to

doubt the wisdom of this approach. It is not only the case that behaviors differ markedly among groups; within group variability is marked as well. Our evidence leads us to recognize two fundamental types of behavioral heterogeneity: between group heterogeneity, which is apparently closely related to group differences in social structure and culture, and within group heterogeneity, which is for the most part unexplained in our study but which is strongly suggestive of the coexistence within groups of distinct dispositions, situationally cued mental models, or other behavior-producing constructs.

Two central problems are raised by our research. First, our work, along with hundreds of other experiments published in the last two decades raises an evolutionary puzzle: What accounts for the success and persistence of behavior that violates the selfish axiom? We do not doubt that selfish motives are both common and essential to understanding human behavior. The challenge is to understand how and why unselfish behaviors and motives could evolve in the face of the material advantages accruing to selfish individuals. We think that long-run evolutionary processes governing the distribution of genes and cultural practices could have resulted in a substantial fraction of each population acting in certain situations (and perhaps generally) to forego material payoffs in order to share with others or to punish unfair actions, as did our experimental subjects. A number of recent models have shown that under conditions that appear to approximate the ancestral environments of human populations, prosocial behavior (carried in either genes or culture) can proliferate (Bowles et al. 2002, Henrich & Boyd 2001, Boyd et al. 2001, Gintis 2000 and Bowles & Hopfensitz 2001). But the evolutionary puzzle posed by the violations of the selfishness axiom on the broad canvas of cultural variation in our sample is far from resolved.

The second question raised by our study is: Why did members of different groups behave so differently? Why is there so much variation between human groups, considering we do not observe this degree of variation among most university students or in other animal species? Addressing this question will require theories that explain why and how different dispositions, different sets of contextual rules, or different modes of information processing spread in different groups and how they are maintained. A central task of any such account is to understand why behavioral patterns appear to covary with economic and social structures in the ways we have observed. Failure to recognize the extent of human diversity and the range of processes that have generated the human mosaic, may lead large sections of social science to an empirically false and culturally limited construction of human nature.

Works Cited

- Barr A. (2002) Kinship, Familiarity, and Trust. In *Cooperation, Reciprocity and Punishment: Experiments from 15 small-scale societies*. Eds. Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. & McElreath, R. Book Manuscript.
- Barry, H., Child, I. L. & Bacon, M. K. (1959) Relation of Child Training to Subsistence Economy. *American Anthropologist* 61: 51-63.
- Bernstein, L. (1992) Opting out of the Legal System: Extralegal Contractual Relations in the Diamond Industry. *Journal of Legal Studies* 21:115–157
- Bolton, G. E. & Ockenfels, A. (1999) A Theory of Equity, Reciprocity and Competition. *American Economic Review* 90:166–94.
- Bowles, S. & Hopfensitz, A. (2001) The Co-Evolution of Individual Behaviors and Social Institutions. *Workshop on Groups, Multi-Level Selection, and Evolutionary Dynamics*: Santa Fe Institute.
- Bowles, S. (1998) Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions. *Journal of Economic Literature* 36: 75–111.
- Boyd, R., Bowles, S., Fehr, E. & Gintis, H. (2002) *The Foundations of Social Reciprocity*. In preparation.
- Boyd, R., Gintis, H., Bowles, S. & Fehr, E. (2001) Inter-Demic Group Selection can lead to the Evolution of Group-Beneficial Punishment Even in Larger Groups. Santa Fe Institute.

- Boyd, R., & Richerson, P.J. (1985) *Culture and the Evolutionary process*. Chicago: University of Chicago Press.
- Breer, P. E. & Locke, E. A. (1965) *Task Experience as a Source of Attitudes*. Homewood, Ill.,: Dorsey Press.
- Burks, S., Carpenter, J. & Verhoogen, E. (2001) High Stakes Bargaining with Non-Students. Santa Fe Institute.
- Camerer, C. (2002) *Behavioral Game Theory: Experiments on Strategic Interaction*, Princeton University Press
- Cameron, L. (1999) Raising the Stakes in the Ultimatum Game: Experimental Evidence from Indonesia. *Economic Inquiry*, 37:47–59.
- Caporael, L., Dawes, R. M., Orbell, J. M., & Van de Kragt, A. J. (1989) Selfishness Examined: Cooperation in the Absence of Egoistic Incentives. *Behavioral and Brain Sciences* 12:683–97.
- Charness, G. & Rabin, M. (1999) Social Preferences: Some Simple Tests and a New Model. University of California, Berkeley.
- Davis, D. D. & Holt, C. A. (1993) *Experimental Economics*. Princeton, NJ: Princeton University Press.
- Falk, A. & Fischbacher, U. (1998) A Theory of Reciprocity. Institute for Empirical Economic Research, University of Zurich.
- Fehr, E. & Gächter, S. (2000). Fairness and Retaliation: The Economics of Reciprocity. *Journal of Economic Perspectives*. 14:159–181

- Fehr, E. & Schmidt, K. M. (1999) A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics*, 114:817–68.
- Gächter, S. & Falk, A. (2001) Reputation or Reciprocity? Consequences for Labour Relation. *Scandinavian Journal of Economics*, In press.
- Gintis, H. (2000) *Game Theory Evolving*. Princeton, NJ: Princeton University Press.
- Harris, J. R. (1999) *The Nurture Assumption: Why children turn out the way they do*. The Free Press, New York
- Hayashi, N., Ostrom, E. Walker, J. M., & Yamagishi, T. (1999) Reciprocity, Trust, and the Sense of Control. *Rationality and Society*, 11:27–46.
- Henrich, J. (2000) Does Culture Matter in Economic Behavior? Ultimatum Game Bargaining among the Machiguenga of the Peruvian Amazon. *American Economic Review*, 90: 973–979.
- Henrich, Joseph & Natalie Smith, N. (2002) Comparative experimental evidence from Machiguenga, Mapuche, Huinca & American populations shows substantial variation among social groups in bargaining and public goods behavior. In: *Cooperation, Reciprocity and Punishment: Experiments from 15 small-scale societies*. Ed. Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. & McElreath, R. Book Manuscript.
- Henrich, J. & Boyd, R. (2000) Why People Punish Defectors. Weak Conformist Transmission can Stabilize the Costly Enforcement of Norms in Cooperative Dilemmas. *Journal of Theoretical Biology*. 208: 79–89

- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2002) *Cooperation, Reciprocity and Punishment: Experiments from 15 small-scale societies*. Book manuscript.
- Henrich, J. & McElreath, R. (In press). Are Peasants Risk Averse. *Current Anthropology*.
- Johnson, A. & Earle, T. (2000) *The Evolution of Human Societies: From Foraging Group to Agrarian State*. Second Edition. Stanford, CA: Stanford University Press.
- Kagel, J. H. & Roth, A.E. (1995) *Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press.
- Kaplan, H. & Hill, K. (1985) Food Sharing among Aché Foragers: Tests of Explanatory Hypotheses. *Current Anthropology*, 26:223–46.
- Kohn, M. (1990) Position in the Class Structure and Psychological Functioning in the U.S., Japan, and Poland. *American Journal of Sociology*, 95:964–1008.
- Kollock, P. (1994) The Emergence of Exchange Structures: An Experimental Study of Uncertainty, Commitment, and Trust. *American Journal of Sociology*, 100:313–45.
- Loewenstein, G. (1999) Experimental Economics from the Vantage Point of View of Behavioural Economics. *Economic Journal*, 109: F25–F34.
- Marlowe, F. (2002) Dictators and ultimatums in an egalitarian society of hunter-gatherers, the Hadza of Tanzania. In *Cooperation, Reciprocity and Punishment: Experiments from 15 small-scale societies*. Ed. Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. & McElreath, R. Book Manuscript
- Ostrom, E. (1998) A Behavioral Approach to the Rational Choice Theory of Collective Action. *American Political Science Review*, 92:1–21.

- Ostrom, E. (2000) Collective Action and the Evolution of Social Norms. *Journal of Economic Perspectives*, 14:137–58.
- Pillutla, M.M. and Chen, X. P. (1999) Social Norms and Cooperation in Social Dilemmas: The Effects of Context and Feedback. *Organizational Behavior & Human Decision Processes*, 78:81–103.
- Plomin, R. (1997) Behavioral genetics (3rd ed.). New York: W.H. Freeman.
- Ross, L. & Nisbett, R. E. (1991) *The Person and the Situation: Perspectives of Social Psychology*. Philadelphia: Temple University Press.
- Roth, A. E. (1995) Bargaining Experiments. In: *The Handbook of Experimental Economics*. Ed. Kagel, J. & Roth, A. Princeton, NJ: Princeton University Press.
- Roth, A. E., Prasnikar, V., Okuno-Fujiwara, M., & Zamir, S. (1991) Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study. *American Economic Review*, 81:1068–95.
- Sherif, M. (1937) An Experimental Approach to the Study of Attitudes. *Sociometry*, 1: 90–98.
- Siamwalla, A. (1978) Farmers and Middlemen: Aspects of Agricultural Marketing in Thailand. *Economic Bulletin for Asia and the Pacific*, 39:38–50.
- Simon, H. A. (1990) A Mechanism for Social Selection and Successful Altruism. *Science*, 250:1665–67.
- Slonim, R. and Roth A. E. (1998). Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic. *Econometrica* 66:569–596.

- Smith, E.A., Bowles, S. & Gintis, H. (2001) Costly Signaling and Cooperation. Santa Fe Institute.
- Smith, N. (2001) Ethnicity, reciprocity, reputation and punishment: an ethnoexperimental study among the Chaldeans and Hmong of Detroit. Unpublished dissertation: University of California, Los Angeles.
- Sober, E. & Wilson, D. S. (1994) Reintroducing Group Selection to the Human Behavioral Sciences. *Behavioral and Brain Sciences* 17:585–654.
- Tversky, A. & Kahneman, D. (1992) Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty* 5: 297–323.
- Woodburn, J. (1968) An Introduction to the Hadza Ecology. In *Man the Hunter*. ed. Lee, R. B. & DeVore, I. Chicago: Aldine, pp. 49–55.

Table 1. Ethnographic Summary of Societies

Group	Language Family	Environment	Economic Base	Residence	Complexity	Researcher	PC	MI
Machiguenga	Arawakan	Tropical Forest	Horticulture	Bilocal semi nomadic	Family	Henrich, Smith	1	4
Quichua	Quichua	Tropical Forest	Horticulture	Sedentary/ Semi-nomadic	Family	Patton	1	2
Achuar	Jivaroan	Tropical Forest	Horticulture	Sedentary/ Semi-nomadic	Family plus extended ties	Patton	5	2
Hadza	Khoisan/Isolate	Savanna-Woodlands	Foraging	Nomadic	Band	Marlowe	4	1
Aché	Tupi-Guarani	Semi-tropical Woodlands	Horticulture/ Foraging	Sedentary- Nomadic	Band	Hill, Gurven	6	4
Tsimane	Macro-Panoan Isolate	Tropical Forest	Horticulture	Semi-nomadic	Family	Gurven	1	3
Au	Toricelli/ Wapei	Mountainous Tropical Forest	Foraging/ Horticulture	Sedentary	Village	Tracer	3	5
Gnau	Toricelli/ Wapei	Mountainous Tropical Forest	Foraging/ Horticulture	Sedentary	Village	Tracer	3	5
Mapuche	Isolate	Temperate Plains	Small scale farming	Sedentary	Family plus extended ties	Henrich	2	6
Torguuds	Mongolian	High latitude desert Seasonally-flooded grassland	Pastoralism	Transhumance	Clan	Gil-White	2	8
Kazakhs	Turkic	High-latitude Desert Seasonally-flooded grassland	Pastoralism	Transhumance	Clan	Gil-White	2	8
Sangu	Bantu	Savanna-Woodlands Seasonally-flooded grassland	Agro-Pastoralists	Sedentary or Nomadic	Clan-Chiefdom	McElreath	5	8
Orma	Cushitic	Savanna-Woodlands	Pastoralism	Sedentary or Nomadic	Multi-Clan Chiefdom	Ensminger	2	9
Lamalera	Malayo-Polynesian	Island Tropical coast	Foraging-Trade	Sedentary	Village	Alvard	7	7
Shona	Niger-Congo	Savanna-Woodlands	farming	Sedentary	Village	Barr	5	8

Table 2 : Ultimatum Game Experiments

Group	Sample Size	Stake	Mean	Mode (% sample) ¹	Rejections	Low rejections ²
Lamalera ³	19	10	0.57	0.50 (63%)	4/20 (sham) ⁴	3/8 (sham)
Aché	51	1	0.48	0.40 (22%)	0/51	0/2
Shona (Resettled)	86	1	0.45	0.50 (69%)	6/86	4/7
Shona (all)	117	1	0.44	0.50 (65%)	9/118	6/13
Orma	56	1	0.44	0.50 (54%)	2/56	0/0
Au	30	1.4	0.43	0.3 (33%)	8/30	1/1
Achuar	14	1	0.43	0.50 (36%)	2/15 ⁵	1/3
Sangu (herders)	20	1	0.42	0.50 (40%)	1/20	1/1
Sangu (farmers)	20	1	0.41	0.50 (35%)	5/20	1/1
Sangu	40	1	.41	0.50 (38%)	6/40	2/2
Shona (Unresettled)	31	1	0.41	0.50 (55%)	3/31	2/6
Hadza (big camp)	26	3	0.40	0.50 (35%)	5/26	4/5
Gnau	25	1.4	0.38	0.4 (32%)	10/25	3/6
Tsimane	70	1.2	0.37	0.5/0.3 (44%)	0/70	0/5
Kazakh	10	8	0.36	0.38 (50%)	0/10	0/1
Torguud	10	8	0.35	0.25 (30%)	1/10	0/0
Mapuche	31	1	0.34	0.50/0.33 (42%)	2/31	2/12
Hadza (all camps)	55	3	0.33	0.20/0.50 (47%)	13/55	9/21
Hadza (small camp)	29	3	0.27	0.20 (38%)	8/29	5/16
Quichua	15	1	0.25	0.25 (47%)	0/14 ⁶	0/3
Machiguenga	21	2.3	0.26	0.15/0.25 (72%)	1	1/10

¹If more than one mode is listed, the first number is the most popular offer and the second number is the second most popular, etc. The percent in parentheses is the total proportion of the sample at the mode(s). For example, for the Machiguenga 72% of the sample offered either 0.15 or 0.25.

²This is the frequency of rejections for offers equal to or less than 20% of the pie.

³In Lamalera, Alvard used pack of cigarettes instead of money to avoid the appearance of gambling. Cigarettes can be exchanged for goods/favors.

⁴Instead of giving responder the actual offers, Alvard gave 20 ‘sham’ offers that range from 10% to 50% (mean sham offer = 30%). These are the frequency of responses to sham offers.

⁵ Patton randomly paired Quichua and Achuar players, and as a result there were 14 Achuar proposers and 15 Achuar responders.

⁶ Patton randomly paired Quichua and Achuar players, and as a result there were 15 Quichua proposers and 14 Quichua responders

Table 3: Summary of Public Good Experiments

Group	Format ¹	Group size	MPCR ²	Sample Size	Stake ³	Mean	Mode ⁴	% Full Cooperation	% Full Defection
Machiguenga	CPR	4	0.375	21	0.58	0.22	0 (38%)	0	38
Tsimané	VC	4	0.50	134	0.75	0.54	0.67 (17%)	1.5	5
Mapuche	VC	5	0.40	12	0.33	0.34	0.1 (42%)	0	0
Huinca	VC	5	0.40	12	0.33	0.58	0.5 (25%)	17	0
Ache	VC	5	0.40	64	1	0.65	0.40 (30%)	3.1	1.6
Orma	VC	4	0.50	24	0.5	0.58	0.40 (37%)	25	0
Michigan	CPR	4	0.375	64	0.58	0.43	0 (33%)	26	33
Swiss⁵	VC	4	0.375	120	0.1	0.33	0 (45%)	14	45
Swiss prttrs	VC	4	0.375	96	0.1	0.55	1 (24%)	24	9.60

¹ Our Public Goods Experiments have two formats with identical payoff structures. In Common Pool Resource (CPR) Games, each player simultaneously withdraws between 0 and some fixed amount from a common pot. Whatever remains in the pot after all the players have withdrawn is increased and distributed equally among all players. In Voluntary Contributions Games (VC), players are each endowed with some amount of money. Players then contribute any amount they want, between zero and their endowment, to a common pot (or a ‘project’). The amount total contributed to the common pot is increased and distributed equally among all players.

² Marginal per capita return

³ Stakes sizes are standardized to one-day wage in the local market, so this column is the endowment received by each player divided by one-day’s wage.

⁴ The percent in parentheses is the total proportion of the sample at the mode.

⁵ The Swiss data comes from Fehr & Gächter (1999). The ‘Swiss’ row represents data from the first five rounds of a ‘strangers treatments’ in which players never played with the same people more than once. We aggregated this data because the individual rounds were indistinguishable from one other. From the same study, the ‘Swiss prttrs’ data is the first round of a 10 round game in which players repeatedly played with the same players through all 10 rounds.

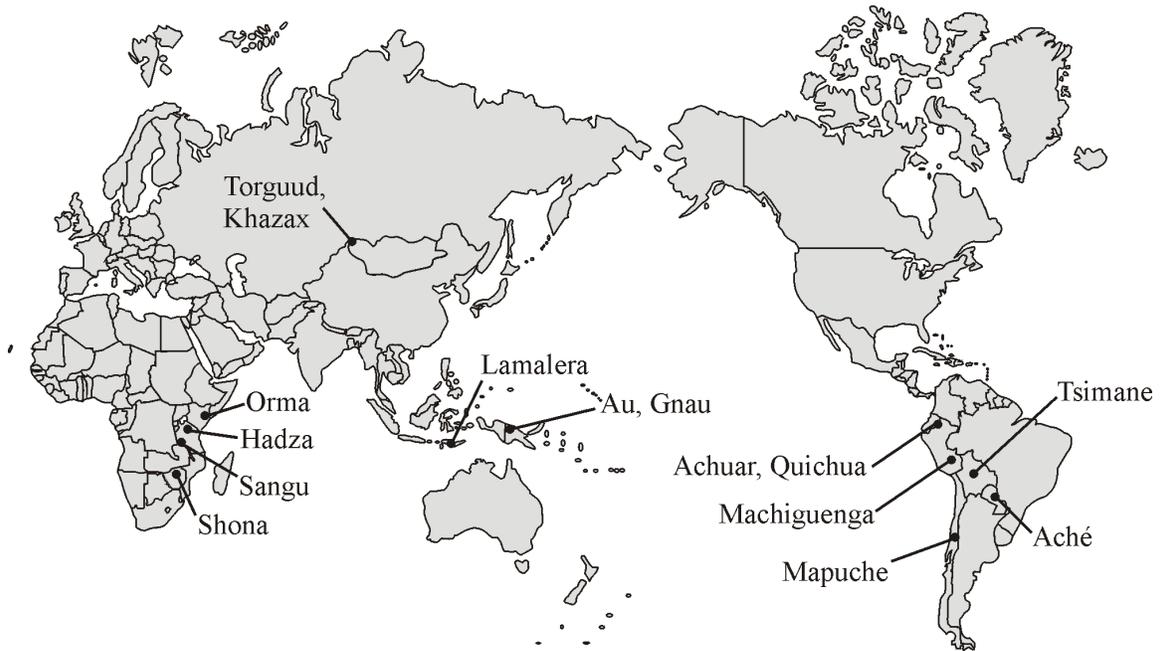


Figure 1. Locations of the societies mentioned in the text.

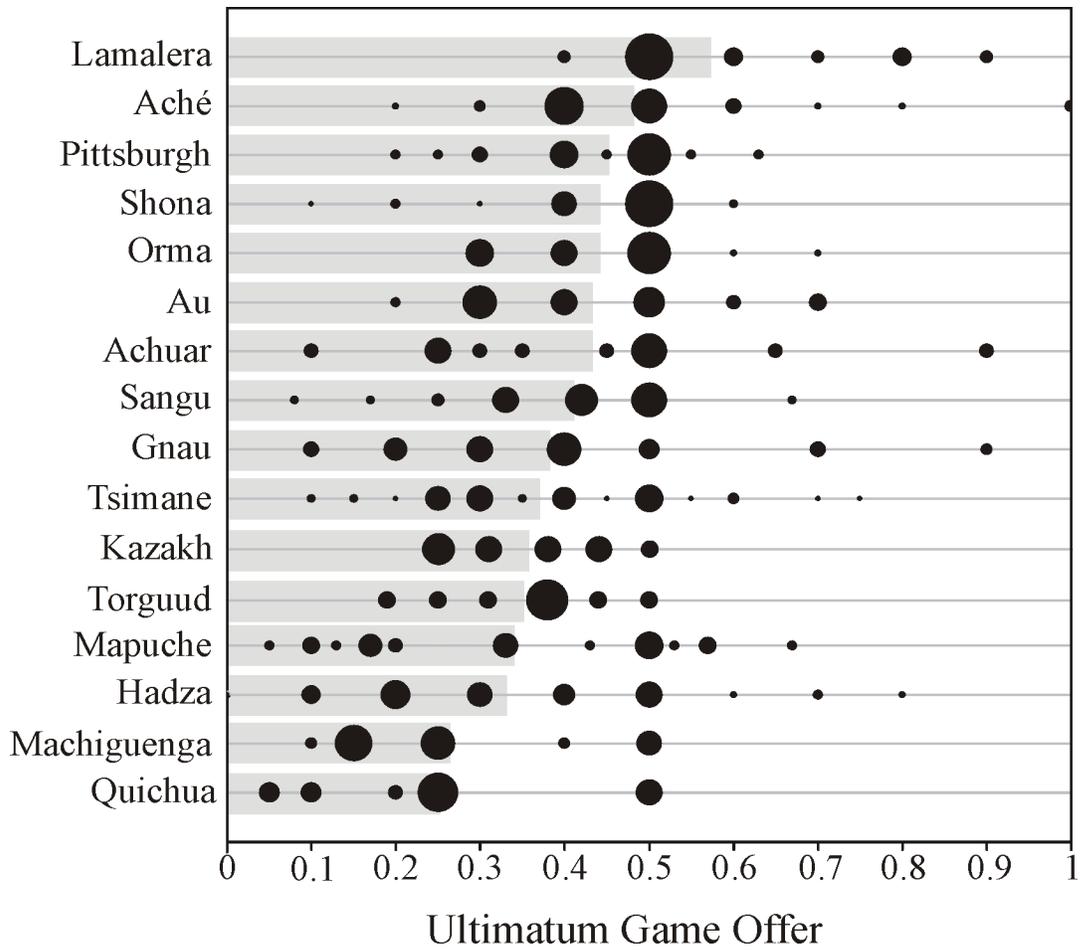


Figure 2. A Bubble Plot showing the distribution of Ultimatum Game offers for each group. The size of the bubble at each location along each row represents the proportion of the sample that made a particular offer. The right edge of the lightly shaded horizontal gray bar is the mean offer for that group. Looking across the Machiguenga row, for example, the mode is 0.15, the secondary mode is 0.25, and the mean is 0.26.

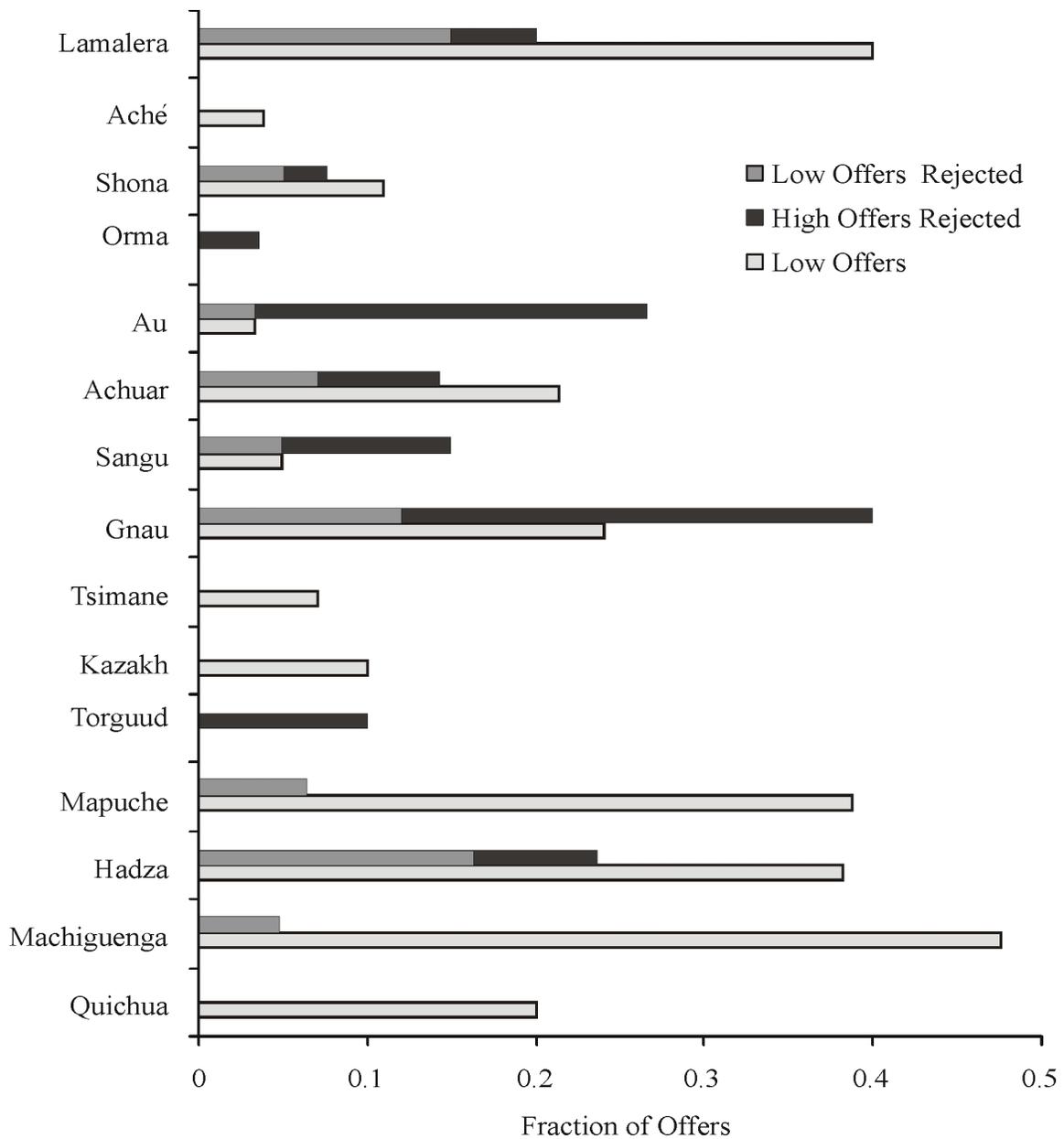


Figure 3. Summary of Ultimatum Game Responder's Behavior. The lightly shaded bar gives the fraction of offers that were less than 20% of the pie. The length of the darker shaded bar gives the fraction of all Ultimatum Game offers that were rejected. The gray part of the darker shaded bar gives the number of these low offers that were rejected as a fraction of all offers. The low offers plotted for the Lamalera were sham offers created by the investigator.

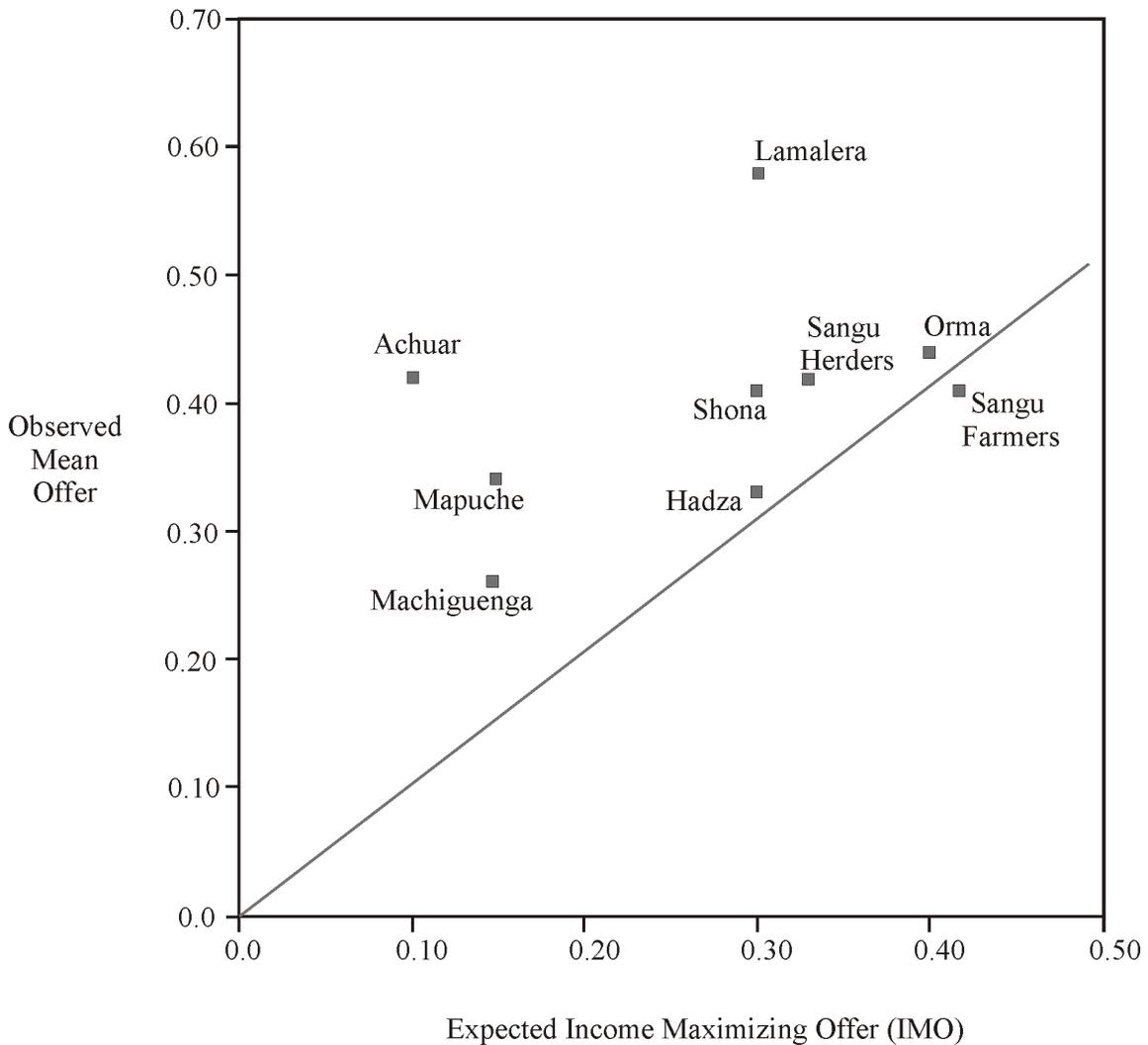


Figure 4. The observed mean Ultimatum Game offer in various groups plotted against the expected income maximizing offer estimated from observed distribution of rejections. All but one of the points lie above the 45 degree line which gives the expected mean offer under the assumption that people are expected income maximizers. We were unable to estimate the IMO for societies with no rejections (Quichua, Tsimane, Ache, Kazakhs), or societies in which rejections bore no systematic relationship to offer (Au, Gnaou, Torguuds).

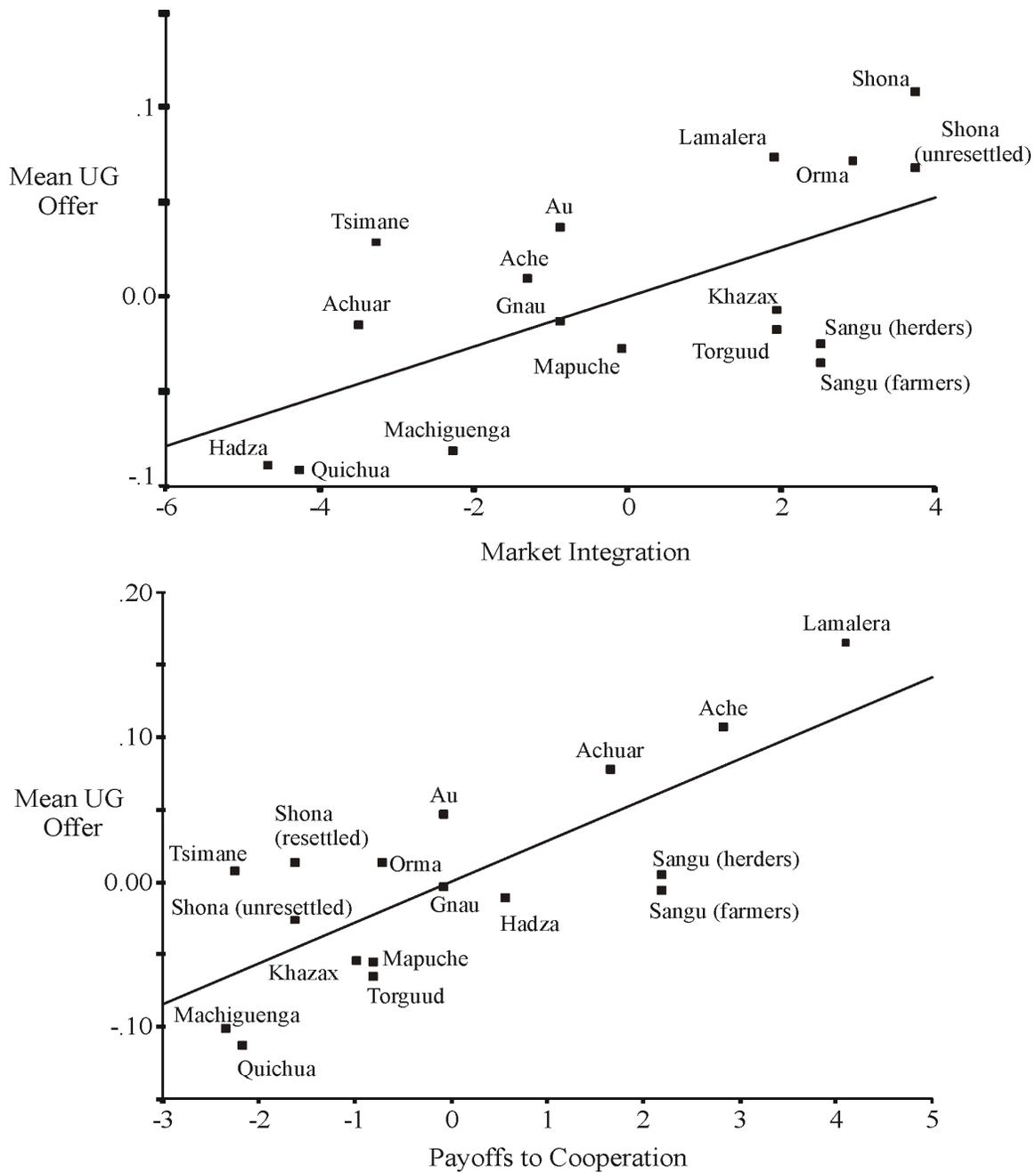


Figure 5. Partial regression plots of mean Ultimatum Game offer as a function of indexes of Market Integration and Payoffs to Cooperation. The vertical and horizontal axes are in units of standard deviation of the sample. Because MI and PC are not strongly correlated, these univariate plots give a good picture of the effect of the factors captured by these indexes on the Ultimatum Game behavior.