The Economic Journal, **113** (*July*), 614–630. © Royal Economic Society 2003. Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

TRUST AND EXPECTED TRUSTWORTHINESS: EXPERIMENTAL EVIDENCE FROM ZIMBABWEAN VILLAGES*

Abigail Barr

An economic experiment involving 24 small, tightly knit communities allows us to distinguish between trusting or trust-like behaviour based on expectational and non-expectational motivations. A model linking trusting behaviour to expectations of trustworthiness explains over half of the variation across communities. However, the estimated parameters are different (while being similarly well defined) for traditional and resettled communities. This is taken as evidence that non-expectational motivations are at work and vary with community type. Both the data and certain stylised facts suggest that altruistic motivations matter less and motivations relating to a desire to 'community-build' matter more in resettled communities.

Trust is valuable. It 'is an important lubricant of a social system. It is extremely efficient; it saves a lot of trouble to have a fair degree of reliance on other people's word' (Arrow, 1974, p. 23). Empirically, trust is associated with higher rates of economic growth, and the emergence and effective functioning of large-scale organisations, including governments (Knack and Keefer, 1997; Fukuyama, 1995; LaPorta *et al.*, 1997). How, then, can we promote trusting behaviour? The answer to this question depends on how trusting behaviour is motivated. If it is motivated by an expectation of trustworthiness, its dynamics are codetermined with those of trustworthiness and its promotion reduces to the promotion of trustworthiness. But if it is also motivated by other factors, its dynamics may be at least partially independent of those of trustworthiness and it may be appropriate to promote trusting behaviour independently or as a means of promoting trustworthiness. Ultimately the two must grow together – what is at issue here is the appropriate point of entry for a trust-promoting intervention.

That trusting behaviour can be motivated by expected trustworthiness is uncontroversial. Indeed, many modern definitions of trust, including those proposed by Gambetta (1988), Yamagishi and Yamagishi (1994) and Hardin (2001), state that it is by nature an expectation. By distinguishing between trust and trusting or trust-like behaviour we introduce the possibility that the latter might also be motivated by other factors. Consider the basic trust game of Bacharach and Gambetta (2001) an example of which is presented in Figure 1. If Agent *R*'s preferences are such that the Raw Payoff Matrix is augmented and thus transformed into All-in Payoff Matrix I, she will behave trustworthily. A transformation of this kind would occur if Agent *R* were altruistic or inequality averse or have a preference for reciprocating kindness (Levine, 1998; Fehr and Schmidt, 1999;

* Many thanks to Michael Shambare for his tireless efforts in the field. Thank you also to my two anonymous referees, to Michael Bacharach and to members of the Cross-Cultural Experimental Group, and the Centre for the Study of African Economies all of whom offered invaluable comments. Finally, I owe a huge debt of thanks to the villagers who took part in the study. All remaining errors are my own. The Department for International Development funded this research through ESCOR grant R7650.

Raw Payoff Matrix

		Agent *R*	
		be trustworthy	be untrustworthy
Agent *T*	behave in trusting manner	5, 2	0, 3
	behave in untrusting manner	2, 0	2, 0

All-in Payoff Matrix I

		Agent *R*	
		be trustworthy	be untrustworthy
Agent *T*	behave in trusting manner	5, 4	0, 3
	behave in untrusting manner	2, 2	2, 0

All-in Payoff Matrix II

		Agent *R*	
		be trustworthy	be untrustworthy
Agent *T*	behave in trusting manner	8, 2	3, 3
	behave in untrusting manner	2, 0	2, 0

Fig. 1. *The Basic Trust Game, Raw and All-in Payoffs*

Bolton and Ockenfels, 2000; Rabin, 1993; Dufwenberg and Kirchsteiger, 2000).[1] If Agent *T* believes that Agent *R* has such preferences, i.e., expects Agent *R* to be trustworthy, it is rational for Agent *T* to behave in a trusting way.

Now suppose that Agent *T*'s preferences are such that the Raw Payoff Matrix is transformed into All-in Payoff Matrix II, then Agent *T* would behave in a trusting manner regardless of her expectations. Yamagishi and Yamagishi (1994) and Hardin (2001) would argue that in this case we are no longer looking at a trust game. However, others identify a trust game with reference to its raw payoff matrix

[1] See Appendix A for the general formulations of the matrices in Figure 1.

only and accept that trusting behaviour may also be motivated by the payoff-transforming preferences of Agent *T*. Sheppard and Tuchinsky (1996), for example, argue that if Agent *T* identifies with the same social group as Agent *R*, she may behave in a trusting way because she believes that it would be for the greater good.[2]

In this paper, I present the results of an economic experiment that allows me to distinguish between trusting or trust-like behaviour based on payoff transformations associated with Agent *R*'s preferences and trusting or trust-like behaviour based on payoff transformations associated with Agent *T*'s preferences, i.e., I distinguish between trusting behaviour based on expectations and trusting behaviour based on two distinct types of shared social identity.

Berg *et al.* (1995) (BDM) found that trusters behave differently in a one-shot investment game if they know the results of earlier investment games. This suggests that expectations are important in determining trusting behaviour. However, the BDM experiment neither occurred within nor induced a social context in which a shared identity could have existed or emerged and so cannot facilitate the comparison I wish to make. I could have adapted the BDM experiment to include a shared-identity-inducing dimension but I chose an alternative approach. I conducted an experiment similar to BDM's within 24 pre-existing, small, tightly knit communities. This approach lends the experiment a verisimilitude that would be lacking in the laboratory, as it facilitates a comparison of the effects of expectations and shared identity when both are formed as a result of real community life.

In each of the 24 community-specific sessions the players did not know the precise identity of their co-player but did know that it was someone from their own community. Hence, while there was no easy means at their disposal to deter untrustworthy behaviour, they would have (a) possessed some knowledge upon which to base their expectations about, and (b) shared a community-specific identity with their co-player.[3]

The 24 communities were all in rural Zimbabwe. 18 were the result of a resettlement exercise, which took place in the early 1980s, while the remaining six were more traditional. The distinction between resettled and traditional communities has important implications for the nature of the social identities that the villagers share with their potential co-players. In the traditional communities social identity is strongly associated with kinship; traditional villages are bound together by dense networks of kinship ties. In the resettled communities, while some kinship ties between households do exist, they are rare.[4] Here, social identities appear to have

[2] If the social context implies that the game is repeated, Agent *R* could be deterred from untrustworthiness by the threat of exclusion from subsequent games and Agent *T* would 'trust' even without payoff-transforming preferences. Weber (1968 [1922]), Shapiro *et al.* (1992), and Sheppard and Tuchinsky (1996) call this deterrence-based trust, while Yamagishi and Yamagishi (1994), Hardin (2001), and Bacharach and Gambetta (2001) argue that, rather than inducing trust, deterrents reduce the need for trust. Within the conceptual framework proposed here, one would say that the deterrents change rather than augment the raw payoff matrix, rendering the game no longer a trust game.

[3] Glaeser *et al.* (1999) introduce real social contexts into a trust experiment. However, their design is such that the primary effect of the social contexts is to facilitate deterrence.

[4] This is because individual applicant households were placed in the new villages with reference to their place on the waiting list and their status relative to a set of predetermined criteria, while no or rare consideration was given to whether related households had also applied and whether they wished to be resettled alongside these relatives.

emerged as a result of the shared experience of resettlement. Of a sample of 372 of villagers drawn from the resettled communities included in this study, 96% stated that the sentence 'Resettlement provided us with a chance to make a new start' applied perfectly to themselves and their families. Their commitment to this new start has manifested itself in several ways. In particular, despite or possibly as a response to the lack of kinship ties, the resettled villagers have formed many clubs and associations. The average resettled household was maintaining the same number of associational memberships as a household in a traditional village after only four years of resettlement, twice as many by the year 2000, and was, even then, becoming increasingly involved in civil society. In their own words, the resettled villagers are working towards 'living in harmony with one another as God (or the ancestors, depending on the speaker) would wish.'

So, we have two distinct bases for shared social identities that are present to different degrees within the two types of village. *Ceteris paribus,* to the extent that the greater good equates with pure altruism and the latter relates to genetic closeness (Hamilton, 1964), we would expect trust-like behaviour based on shared social identity to be more commonplace in traditional communities. However, to the extent that the greater good equates with the promotion of trusting and trustworthy behaviour and investment in the longer term functioning and prosperity of the community, we would expect trust-like behaviour to be more commonplace in resettled communities.

The results of the experiment suggest that trusting behaviour is partially based on expectations about people's trustworthiness. In communities where trustworthiness was low, players behaved in a less trusting manner. In addition, consistent with the players being risk averse, in communities where there was a higher variance in trustworthiness, players behaved in a less trusting manner. This notwithstanding, in the 18 resettled communities trusting behaviour was less responsive to variations in trustworthiness than in the six traditional communities. This is consistent with the theory that the resettled villagers are trying to community-build and have partially compensated for lower levels of kin-based altruism.

The paper has five Sections. Following this introduction, in Section 1 I describe the experimental protocol. In Section 2, I describe the framework used to analyse the data. In Section 3, I present the data, provide a comparison of my results with those of BDM, and then turn to the detailed analysis of the Zimbabwean villagers' behaviour. Finally, in Section 4 I conclude with a discussion about the results and the insights they provide.

1. Experimental Protocol

BDM's investment game is a one shot game played by pairs of subjects. Both players are given an initial endowment of money. The first player has the option of giving some of his/her money to the second player, i.e., of making an investment. Whatever he/she chooses to invest is tripled before being given to the second player. The second player then has the option of giving some portion of the tripled amount back to the first player. The one-shot nature of the game combined with player anonymity, ensures that there is no possibility that reputation mechanisms

based on repeated interactions, contractual pre-commitments or threats of punishment are generating the results. It ensures that first players who chose to invest more than zero are vulnerable to exploitation by the second player. The more first players invest, the greater their potential return, but the greater their vulnerability.

In my version of the game, each player's initial endowment was 20 Zimbabwean dollars and all play was conducted in Zim$5 notes.[5] Thus, the first player had to choose $s \in S$, where $S = \{0, 5, 10, 15, 20\}$. The choice of s by the first player determined the $\Gamma(s)$ sub game, in which the second player chose $r \in R$, where $R(s) = \{0, 5, 10, 15,\ldots,3s\}$. This game was played once in each of 24 sessions. Each session was held in a different Zimbabwean village. Each player was paired with someone from his or her own village and informed accordingly. In each village there were between 4 and 11 pairs of volunteers playing the BDM game. A total of 141 pairs played this game. In every session there were other volunteers present who were playing an ultimatum game instead. Each volunteer came from a different household and attended the experimental session in their own village.[6] Just under half of the volunteers (47%) were female and just over half (52%) were heads of households. The allocation of partners, games and roles was determined randomly prior to the start of each session.

Because of the relatively low levels of education and potentially high incidences of illiteracy, experimenters working in developing countries adopt procedures that deviate significantly from those typically applied in Europe and the US. I adopted procedures similar to Henrich (2000). During a pilot exercise a script for the game was developed in Shona, the local language in all the communities.[7] The script contained three components: a detailed and repetitive description of the game; a set of examples and questions showing how particular combinations of decisions yield particular pay-offs for each player; and, for second players only, a description of what their corresponding first players had decided to do. One set of examples was used for all first players and another for all second players. These were designed to demonstrate the key features of the game, while minimising the extent to which players might be led to behave in a certain way. Once the pilot was over the script was closely adhered to. When players had questions, the relevant part of the script was repeated. Players who could not demonstrate that they understood were not allowed to play. The need to test players' understanding verbally rendered a double blind procedure impractical. Hence, potential subject-experimenter effects must be considered.[8] To minimise the impact of such effects on the comparative

[5] This led to average earnings of approximately half a day's casual wage. The exchange rate at the time of the sessions was in the region of 50 Zim $ per UK pound.

[6] The households from which the players originate are participants in an on-going assessment of the effects of resettlement in Zimbabwe. Each household in the assessment exercise was asked to send one adult (above the age of 14) volunteer to the session in their village. In 13 of the communities the assessment and, hence, the experiment involve all households. In the remaining 11, a random sample of households is involved. The chairman or headman of each village was charged with the duty of ensuring that the volunteers arrived at the session. They were told that the volunteers would be playing games and could win some money.

[7] Data from the 16 pairs of players in the pilot have been included throughout the analysis. Excluding these observations does not alter any of the conclusions.

[8] Hoffman *et al.* (1994) provide evidence that subject/experimenter anonymity affects behavior. However, Roth (1995) reviews several other studies that report the opposite.

results, great care was taken to follow the same procedure with each player. This notwithstanding, different subject-experimenter effects on male and female players could have biased the results.[9] The proportion of women volunteers varied from 14.3% to 81.3% across village sessions. I used econometric methods, *ex post,* to test and control for this possible source of bias.

In each session, the players of both games were gathered together. One-by-one they were called to meet with the experimenter and a Shona-speaking research assistant (RA) in private to learn and play the game. In order to avoid player contamination the players were told nothing about the games prior to their individual meetings with the experimenter and RA. Within communities, those who had already played were not allowed contact with those who were still waiting. Where there was a risk of contamination between communities, sessions were held on the same or consecutive days. Interviews with players and other villagers indicated that between-village contamination did not occur. Finally, in the communities where there were relatively few players, their sense of anonymity might have been partially compromised.[10] While no experimental solution could be applied, econometric methods were used, *ex post,* to test and control for this possibility.

2. Analytical Framework

If trusting behaviour is based on expectations, we should be able to model first players' behaviour as an investment decision under uncertainty. Under standard assumptions, a selfish expected utility maximiser's investment would depend positively on his expected return and, if he is risk averse, negatively on the perceived variance in that return. Letting the random variable $\tilde{r}$ with a probability distribution defined on $R(s)$ represent a first player's expectations of the amount that will be returned by the second player in sub game $\Gamma(s)$, the linear approximation to the first player's investment function defined on $\tilde{r}$ would have the following form:

$$s = \alpha_0 + \alpha_1 \mathrm{E}(\tilde{r}) + \alpha_2 \sigma_{\tilde{r}}^2 \tag{1}$$

with α_1 positive and, α_2 negative.

I do not have data on each first player's expectation and perceived variance of $\tilde{r}$. However, the design of the experiment is such that the actual responses of the second players can be used to construct proxies for both $\mathrm{E}(\tilde{r})$ and $\sigma_{\tilde{r}}^2$. This approach is potentially problematic as the set of possible responses from which each second player chooses is determined by the choice made by the corresponding first player. Thus, using second players' actual responses to construct an explanatory variable to include in the analysis of first players' behaviour could lead to problems of reverse causality and endogeneity. An alternative base for these two explanatory variables is the proportional response, i.e., the ratio r/s. This is attractive for two reasons. First, certain values of the proportional response lend themselves to easy interpretation. A proportional response of zero corresponds to self-interested

[9] Eagly and Crowley (1986) found that subject-experimenter effects can differ for men and women.
[10] Hoffman *et al.* (1996) show that reduced subject/subject anonymity leads to greater observed generosity.

money maximisation, a proportional response of one corresponds to what one might call 'pure reciprocity', i.e., returning exactly what was given: and a proportional response of two corresponds to what one might call 'pure sharing', i.e., to dividing the total money in the game equally.[11] These focal points might be as attractive to the players as they are to the author and so assume an important role in determining their behaviour and expectations. Second, the ratio r/s is bound by zero and three whatever the value of s. This reduces but does not eliminate the possibility of a reverse causal link – a second player might give a higher proportional response to a more trusting first player. To eliminate this possibility, I conduct an analysis of second player behaviour, estimating both:

$$r = b_0 + b_1 s + e_1 \tag{2a}$$

and

$$r/s = c_0 + c_1 s + e_2 \tag{2b}$$

where the error terms, e_1 and e_2, are assumed to be i.i.d. normal, before moving on to the analysis of first player behaviour. As long as r/s is found not to depend on s, this analysis can be based on an estimation of the following:

$$\bar{s} = a_0 + a_1\overline{(r/s)} + a_2\mathrm{var}(r/s) + e_3 \tag{3}$$

where $\bar{s}$ is the community mean first players' choice, $\overline{(r/s)}$ is the community mean proportional response by second players, $\mathrm{var}(r/s)$ is the community-level variance in the proportional responses, and e_3 is assumed to be i.i.d. normal. This analysis must be conducted at the community level because of the way in which the explanatory variables are constructed. If a_1 is significantly greater than zero, it lends support to the hypothesis that the first players' trusting behaviour is based on an expectation. If a_2 is significantly less than zero, it lends further support to this hypothesis, while also indicating that the players are risk averse.

When estimating the equations presented above, I include two additional right hand side variables. *Female,* a dummy that takes the value one for female players and zero otherwise, is included to control for possible variations in experimenter effects between male and female players. *Session,* the number of volunteers present in the session that the player attended, is included to control for any reduction in perceived anonymity associated with attending a session involving fewer players.

Initially, the model of trusting behaviour is estimated taking no account of potential differences in motivations between the two types of community. This allows us to test the null hypothesis that trusting behaviour is unresponsive to expectations of trustworthiness against the alternative that it is responsive. Then the model is adapted to facilitate a comparison between the two types of community. Here, the null hypothesis that trusting behaviour is similarly and equally responsive to expected trustworthiness in the two types of community, is tested against the alternative is that it is not. If well defined but distinct relationships can

[11] After the first player chooses s, she has $20 - s$ and the second player has $20 + 3s$ making the total money in the game $40 + 2s$. If the second player chooses to return $r = 2s$ (a proportional response of 2), she is left with $20 + s$ and the first player ends up with $20 - s + 2s = 20 + s$.

be identified for resettled and traditional communities it can be taken as evidence that any unexplained variance in the basic model, rather that being due purely to noise in the data, is due in part to variations in non-expectational motivations for trusting behaviour across communities.

The final step in the analysis aims to distinguish between the motivating effect of social identity based on kinship and that based on a desire to community-build. Consider, first, the likely effect of pure, kin-based altruism. *Ceteris paribus* one would not expect a pure altruist to favour an untrustworthy person any more or less than a trustworthy one. Thus, if line *a* in Figure 2 depicts the relationship between trusting behaviour and expected trustworthiness in the absence of altruism, the relationship in the presence of altruism would be depicted by a line like *b,* and greater altruism would be associated with a larger upward shift in the relationship. In contrast to altruism, a preference for promoting trusting and trustworthy behaviour will be revealed only when both are rare. So in this case, the relationship between trusting behaviour and expected trustworthiness would be represented by a line like *c,* and the stronger the preference the more extreme would be the upward pivot of the relationship around its north-east end. By introducing a community-type dummy and the appropriate interaction terms into the regression analysis, we can explore the relative importance of altruism and a preference to promote trust in traditional and resettled communities.

3. Results

Figure 3 shows the distribution of investments made by the 141 first players. There is one modal investment at Zim$10, half of the initial endowment. Figure 4 shows the distribution of proportional responses made by the 129 second players who

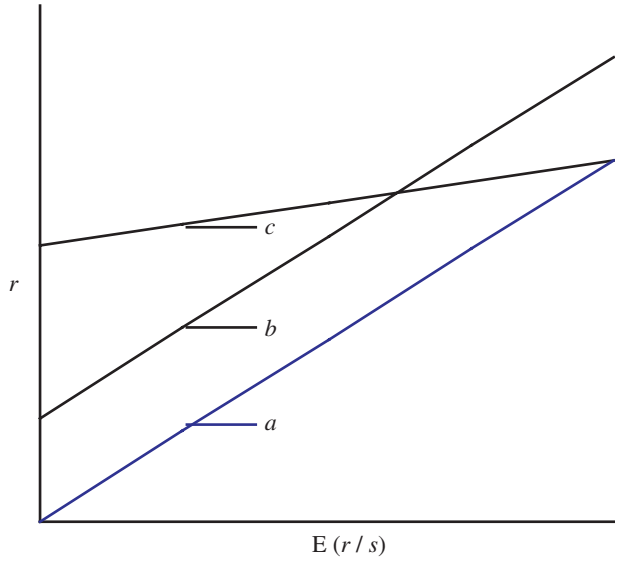

Fig. 2. *Non-expectational Motivations for Trusting Behaviour*

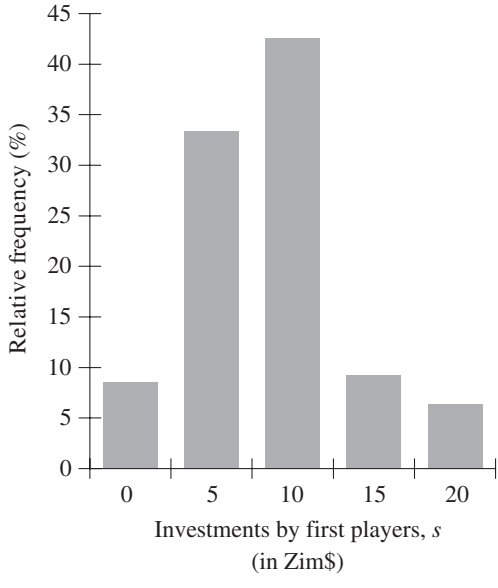

Fig. 3. *Investments by First Players (s)*

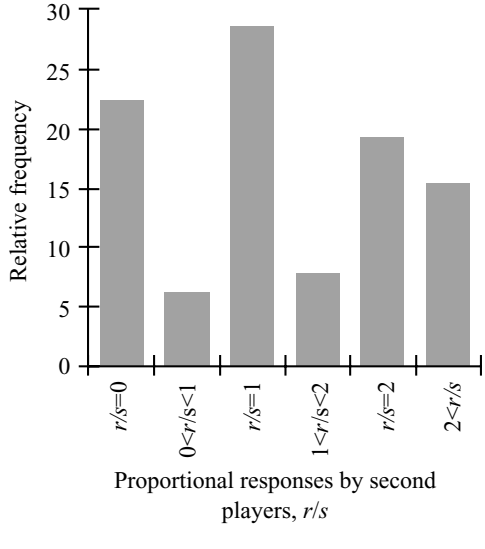

Fig. 4. *Responses by Second Players (r/s)*

received positive, tripled investments. There are three modes at zero, one, and two, indicating that these values did indeed serve as focal points for the players, although whether this is because these values are associated with behavioural norms remains to be seen.

3.1. *Comparison with BDM*

Table 1 contains comparable summary statistics for the Zimbabwean games and those conducted in the US by BDM. Due to the impracticality of a double blind procedure in Zimbabwe, we must exercise caution when making the comparison. This notwithstanding, the comparison serves as a useful check on whether efforts to make the game comprehensible to Zimbabwean villagers, who rarely act anonymously or face situations in which payoffs depend on abstract mathematics, succeeded.

BDM observed few first players investing zero (6%). A marginally greater proportion (9%) of the Zimbabwean first players invested nothing. Further, the mean investment by Zimbabwean first players was lower as a proportion of their initial endowment than that observed in the US experiment (43% as opposed to 52%). On average, second players in the BDM experiment returned less than the first players invested (89%), while in Zimbabwe they returned more (128%). That the Zimbabwean games were not double blind may account for these differences.

Finally, the tri-modal distribution observed in Figure 4 is not present in the BDM data, which display a mode at zero only. This distinction casts some doubt on the hypothesis that the tri-modal distribution in Zimbabwe is the result of 'competition' between behavioural norms. The tri-modal distribution in Zimbabwe could be due to the measures taken to ensure that players understood the game. While efforts were made to minimise the extent to which we biased players' responses to any one of the focal responses, they did feature heavily in the examples because they involved easier mathematics. This notwithstanding, the comparability of the BDM and Zimbabwean data gives us no reason to doubt the understanding of the Zimbabwean players.

3.2. *The Trustworthiness of the Second Players*

Table 2 contains the results of a series of regressions based on the responses of the second players who received positive tripled investments from their respective first players. The regressions in the first and second columns take r, the amount of money returned by the second player, as the dependent variable. The regressions in the third and fourth columns take r/s, the proportional response, as the dependent variable. The regressions in the first and third columns correspond to

Table 1

Comparison Investment Games conducted by Berg et al. (1995) and in Zimbabwe

	Berg *et al.* (1995) (no social history)	Zimbabwe
Number of playing pairs*	32	141
Initial endowment size	10.00 US$	20.00 Zim$
Proportion of first players who invested zero	0.06	0.09
Mean investment by first players	5.16 US$	8.58 Zim$
Mean investment as a proportion of stake	0.52	0.43
Mean response (expressed as a proportion of investment)	0.89	1.28

*Responses are made by second players only when first players invest more than zero.

Table 2

Analysis of Second Players' Responses

	1		2		3		4	
	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
Dependent variable	*r*		*r*		*r/s*		*r/s*	
Constant	4.957	2.333*	5.403	4.997	1.866	0.328***	2.038	0.058*
s	1.270	0.173***	1.026	0.211***	0.004	0.018	−0.025	0.023
female	−3.384	1.462**	−3.370	1.559**	−0.352	0.177**	−0.312	0.188*
session	−0.155	0.074**			−0.021	0.010**		
Joint sig. of community dummies	–		0.0001		–		0.0001	
Observations	129		129		129		129	
R^2	0.325		0.492		0.055		0.301	

Notes: All standard errors have been corrected using White's (1980) procedure. *** significant at 0.01 level, ** significant at 0.05 level, * significant at 0.1 level.

(2*a*) and (2*b*) augmented by *female* and *session*. In the second and fourth columns *session* is replaced with a set of community dummies in order to control fully for any community-level variables. The coefficient on *s* is significant (0.01 level) in the regressions that take *r* as their dependent variable and insignificant (0.1 level) in the regressions that take *r/s* as their dependent variable. Controlling for potentially non-independent error terms within communities does not significantly alter the results. Introducing the square of *s* as an explanatory variable neither alters the results nor improves the fit of the models.

The negative and significant coefficient on *female* might be capturing a differential experimenter effect. However, in informal group discussions the villagers favoured an explanation more akin to a differential cash income effect. They explained that women generally have less control over cash in their households than men. Thus, the female second players, having been presented with their initial endowment plus the tripled amount from the first player, were more reluctant to part with it. The women said that this was especially true if they believed that they were paired with a man because they expected men to spend all their winnings on beer. The negative coefficient on *session* becomes insignificant once I control for potentially non-independent error terms within communities. This notwithstanding, its significance in the OLS regressions suggests that there may have been less perceived anonymity in the smaller sessions. The joint significance (0.0001 level) of the community dummies could be taken as evidence that behavioural rules relating to trustworthiness vary between communities, although it may also be capturing the effects of other community-level variables. That there is significant variation in second player behaviour across communities is an important prerequisite for the community-level analysis of first player behaviour.

3.3. *The Investment Behaviour of First Players*

Given no evidence that *r/s* depends on *s*, I can use *r/s* to derive explanatory variables for the community-level analysis of first players' behaviour. Table 3 contains the results of a series of regressions based on the investments made by the

Table 3

Analysis of First Players' Investments

	1		2		3		4	
	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
Constant	4.077	1.921**	3.892	2.183*	6.554	2.787**	−2.224	2.462
$\overline{(r/s)}$	–		1.941	0.901**	2.009	0.845**	7.021	1.984***
var(r/s)	–		−2.640	1.492*	−2.456	1.376*	1.930	2.040
female	−0.128	0.884	–		–		–	
femalep	–		3.993	2.534	2.431	2.441	0.187	2.375
session	–		0.135	0.067*	0.109	0.064	0.100	0.072
resettled	–		–		−2.192	1.167*	8.267	2.888**
$\overline{(r/s)}$ × *resettled*	–		–		–		−5.325	1.981**
var(r/s) × *resettled*	–		–		–		−4.482	2.623
Joint sig of community dummies	0.0003		–		–		–	
Observations	141		24		24		24	
R^2	0.349		0.331		0.429		0.528	

Notes: All standard errors have been corrected using White's (1980) procedure. *** significant at 0.01 level, ** significant at 0.05 level, * significant at 0.1 level.

141 first players. Before turning to the estimation of (3), consider the individual-level regression in the first column of Table 3. Here, *female* and a set of community dummies are the only explanatory variables. The dummies are jointly highly significant (0.001 level), while the coefficient on *female* is insignificant. The apparent irrelevance of our only individual-level variable combined with the significance of between-community variations indicates that little information is lost in the move to the community-level analysis.

The regression in the second column of Table 3 corresponds to (3). The coefficient on $\overline{(r/s)}$ is positive and significant (0.05 level) and the coefficient on var(r/s) is negative and significant (0.1 level). The significance of the coefficient on *session* does not survive the inclusion of additional variables in the model and so will not be discussed. These results support the hypothesis that first players' trusting behaviour is motivated, at least in part, by expectations of trustworthiness. However, the model explains only one third of the community-level variation in trusting behaviour and the analysis provides no insights into the origin of the other two thirds of the variation.

3.4. *Variations in the Relationship between Trusting Behaviour and the Distribution in Trustworthiness between Resettled and Traditional Communities*

Table 4 presents a comparison of first and second players' behaviour between traditional and resettled communities. The upper panel shows that resettled first players invest significantly (0.05 level) less than their traditional counterparts. For this finding to be consistent with no difference in the relationship between trust and the distribution of trustworthiness across community types, there must be a lower mean level of trustworthiness or a higher variance in trustworthiness in resettled communities. But the statistics in the bottom panel of Table 4 indicate that the distribution of second player's proportional responses is identical across the two types of community.

Table 4

Comparison of First Player Behaviour in Traditional and Resettled Communities

	Traditional		Resettled
Number of playing pairs	32		109
Mean investment (Zim$)	10.47		8.03
Standard deviation (sd) of investments	5.44		4.57
Levene's test for equality of sd (p-value)		0.504	
t-test for equality of mean, equal sd assumed (p-value)		0.012	
Epps-Singleton test for equality of distribution (p-value)		0.000	
Mean proportional response	1.28		1.28
sd of proportional response	0.96		1.04
Levene's test for equality of sd (p-value)		0.698	
t-test for equality of mean, equal sd assumed (p-value)		0.983	
Epps-Singleton test for equality of distribution (p-value)		0.906	

Note: The Epps-Singleton (1986) test is a very powerful non-parametric test that is based on the difference between the characteristic functions of the two samples being compared.

This inconsistency is confirmed by the regressions presented in the third and fourth columns of Table 3. The model in column 3 includes *resettled,* a dummy variable that takes the value one for resettled communities and zero otherwise. The coefficient on *resettled* is negative and significant suggesting that, on average, after controlling for expectational effects, resettled villagers behave in a less trusting way.

The regression in the fourth column of Table 3 contains interaction terms between *resettled* and $\overline{(r/s)}$ and, for completeness, *resettled* and var(r/s). The coefficient on the second interaction term is insignificant, while the coefficient on the first is negative and significant (0.05 level) indicating that trust is significantly less responsive to mean trustworthiness in resettled communities. The coefficient on $\overline{(r/s)}$ now captures the effect of variations in mean trustworthiness on trusting behaviour in traditional communities only. The coefficient tells us that an increase in the mean proportional response of second players from 0.5 to 1.5 is associated with a Zim$7 increase in the mean first player investment. To establish how mean trustworthiness affects trusting behaviour in resettled communities we need to add the coefficients on $\overline{(r/s)}$ and the interaction between *resettled* and $\overline{(r/s)}$ together. This yields a coefficient for resettled communities of 2.539, which is also significantly different from zero (0.1 level). So, in resettled communities an increase in the mean proportional response of second players from 0.5 to 1.5 is associated with only a Zim$2.5 increase in the mean first player investment. The inclusion of the interaction terms changes the sign on the *resettled* dummy. It is now positive and significant indicating that, in the event of both expected trustworthiness and the variance in trustworthiness being zero, resettled villagers would behave in a more trusting manner than traditional villagers. To sum up, in resettled villages trusting behaviour is less responsive to expected trustworthiness than in traditional villages. Thus, while resettled villagers are, on average, less trusting than traditional villagers, there are some resettled villages within which first players make relatively high investments despite low levels of expected trustworthiness.

Holding everything else constant, I plot the relationships between trusting behaviour and expected trustworthiness for the two types of community in

Figure 5. The bold dotted line shows the relationship observed in traditional communities. The bold solid line shows the relationship observed in resettled communities. The shift from the former to the latter can be broken down into two parts, a pivot upward and a shift downwards.

4. Summary and Discussion

The analysis suggests that trusting behaviour is motivated by expectations of trustworthiness. In communities where first players have greater reason to expect second players to be trustworthy, they invest more. Further, consistent with the joint hypothesis that first players are risk averse and trust is an expectation, in communities where first players might reasonably perceive a higher variance in second players' responses, they invest less.

The analysis also indicates that expected trustworthiness is not the only motivation for trusting or trust-like behaviour. The relationship between the two is well defined but distinct for the two types of community involved in the experiment, suggesting that there are other motivations at work that vary with community type. The shift in the relationship between trust and trustworthiness can be broken down into two parts, a downward shift and an upward pivot about the northeast end. The downward shift is consistent with pure altruism being less of a motivation in resettled communities, while the upward pivot is consistent with the stronger preference for community building being more of a motivation in resettled communities.

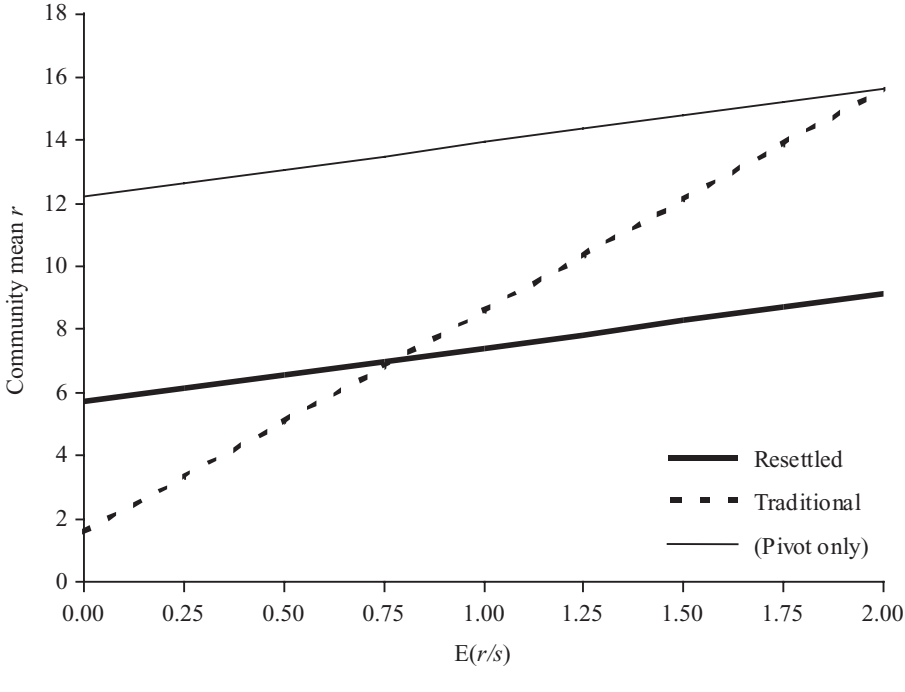

Fig. 5. *Responsiveness of Trusting Behaviour to Expected Trustworthiness in Traditional and Resettled Communities*

While these findings make sense given what we know about kinship, civil social activity and perceptions about the resettlement experience in the communities, it is puzzling that the lower levels of kin-based altruism and greater desire to community build are not manifest in the resettled second players' behaviour as well. If they were, we would not be able to identify the distinct relationships between trust and trustworthiness in the two types of community because trusting behaviour and trustworthiness would increase and decrease in unison. However, this does not constitute an answer to the conundrum.

One possible explanation relating to the effects of community building derives from the asymmetries in the BDM game. First players are invited to look forward, to think about their fellow villagers, to predict and project, and as BDM's name for the game suggests, make decisions about investments that might yield a future reward. Thus, first players may liken the game to their other community building activities. In many of the debriefings this is precisely what they did. In contrast, second players are required to react to what one of their fellow villagers has done rather than predict what any one of them might do. There is nothing in the decision they face that would lead them to look into the future, think of their action as an investment, or consider their community as a whole. They are making a decision about sharing, conditional on a prior action by the fellow villager with whom they are sharing. The debriefings after the BDM games yielded little of relevance here. However, those after the Ultimatum games suggested that when making decisions about sharing, the villagers refer to rules or behavioural norms that they associate with their hunting and gathering ancestors. That second player behaviour is rule or norm driven is also suggested by the tri-modal distribution. That the sharing rules to which villagers adhere pre-date their resettlement is consistent with the marked similarity between the distributions of second player behaviour in resettled and traditional communities and might explain why we see no evidence of the decline in kin-based altruism in the resettled second players' behaviour.

If we accept this explanation for the divergence between first and second player behaviour in resettled villages, another question emerges – are resettled villages in equilibrium? Given that civil social activity is still increasing and the resettled villagers are still talking about community building, probably not. Will community builders continue to invest if, like some of the first players in the trust game, they suffer financial losses as a result? Here, once again it is worth considering the particular characteristics of the BDM game before drawing any conclusions. That the second players do not look forward is an important feature of the game. It derives from its one-shot nature, which is the key to eliminating reputation effects. It is this feature of the game that contrasts most strikingly with the games that the Zimbabwean villagers play during their everyday lives. In everyday life, the game is more likely to involve repeated play with rotating roles. Thus, all the players may be forward-looking and so similarly motivated by the desire to community build.

So, the appropriate point of entry for a trust-promoting intervention would seem to depend on the type of interactions we have in mind. If trustworthiness is low and unresponsive to trust, to promote trusting behaviour in the context of one-shot interactions would increase the vulnerability of the potential trusters. However, where interactions are to be repeated while roles rotate, there may be a

case for encouraging first movers to give others the benefit of the doubt in much the same way as the Zimbabwean resettlers are already doing.

These conclusions are tentative. They are based on a sample of only 24 community-level data points and a rather indirect approach to the identification of the non-expectational motivations for trusting behaviour. This notwithstanding, the research yields new insights into both the motivations for trusting or trust-like behaviour and the BDM game.

University of Oxford

Date of receipt of first submission: May 2000
Date of receipt of final typescript: November 2002

Appendix A

Generalised Payoff Matrices for the Basic Trust Game

Raw Payoff Matrix

		Agent R	
		be trustworthy	be untrustworthy
Agent T	behave trustingly	a_T, a_R	b_T, b_R
	behave untrustingly	c_T, c_R	d_T, d_R

Note: $c_R > a_R$ not necessary but often the case.
Where: $b_T < a_T$, $c_T < a_T$, $d_T > b_T$, $b_R > a_R$, $c_R < a_R$, $c_T = d_T$, $c_R = d_R$.

All-in Payoff Matrix I

		Agent R	
		be trustworthy	be untrustworthy
Agent T	behave trustingly	a_T, a_R^*	b_T, b_R^*
	behave untrustingly	c_T, c_R^*	d_T, d_R^*

e.g., if $a_R^* = a_R + v$, $b_R^* = b_R - v$ and $v > 0.5(b_R - a_R)$.
Note: $c_R^* < a_R^*$ not necessary but often the case.
Where: $b_R^* > a_R^*$, $c_R^* < a_R^*$, $c_R^* = d_R^*$.

All-in Payoff Matrix II

		Agent R	
		be trustworthy	be untrustworthy
Agent T	behave trustingly	a_T^*, a_R	b_T^*, b_R
	behave untrustingly	c_T^*, c_R	d_T^*, d_R

e.g., if $a_T^* = a_T + u$, $b_T^* = b_T + u$, $c_T^* = c_T - u$, $d_T^* = d_T - u$, and $u > 0.5(d_T - b_T)$.
Where: $b_T^* < a_T^*$, $c_T^* < a_T^*$, $d_T^* < b_T^*$, $c_T^* = d_T^*$.

References

Arrow, K. J. (1974). *The Limits of Organization*, New York: W. W. Norton and Co.

Bacharach, M. and Gambetta, D. (2001). 'Trust in signs', in (K. Cook, ed.), *Trust in Society*, New York: Russell Sage Foundation.

Berg, J., Dickhaut, J. and McCabe, K. (1995). 'Trust, reciprocity, and social history', *Games and Economic Behavior*, vol. 10, pp. 122–42.

Bolton, G. E. and Ockenfels, A. (2000). 'A theory of equity, reciprocity, and competition', *American Economic Review*, vol. 90(1), pp. 166–93.

Dufwenberg, M. and Kirchsteiger, G. (2000). 'Reciprocity and wage undercutting', *European Economic Review*, vol. 44, pp. 1069–78.

Eagly, A. H. and Crowley, M. (1986). 'Gender and helping behavior: a meta-analytic review of the social psychological literature', *Psychological Bulletin*, vol. 100, pp. 283–308.

Epps, T. W. and Singleton, K. (1986). 'An omnibus test for the two sample problem using the empirical characteristic function', *Journal of Statistical Computation and Simulation*, vol. 26, pp. 177–203.

Fehr, E., Schmidt, K. M. (1999). 'A theory of fairness, competition, and cooperation', *Quarterly Journal of Economics*, vol. 114(3), pp. 817–68.

Fukuyama, F. (1995). *Trust*, New York: Free Press.

Gambetta, D. (1988). 'Can we trust trust?', in (D. Gambetta, ed.), *Trust: Making and Breaking Cooperative Relations*, Oxford: Basil Blackwell.

Glaeser, E., Laibson, D., Scheinkman, J. and Soutter, C. (1999). 'What is social capital? The determinants of trust and trustworthiness'. NBER working paper no. 7216.

Hamilton, W. D. (1964). 'The genetical evolution of social behaviour', *Journal of Theoretical Biology*, vol. 7, pp. 17–52.

Hardin, R. (2001). 'Trusting persons, trusting institutions', in (R. J. Zeckhauser, ed.), *Strategy and Choice*, Cambridge MA: MIT Press.

Henrich, J. (2000). 'Does culture matter in economic behaviour? Ultimatum game bargaining among the Machiguenga of the Peruvian Amazon', *American Economic Review*, vol. 90(4), pp. 973–9.

Hoffman, E., McCabe, K., Shachat, K. and Smith, V. (1994). 'Preferences, property rights and anonymity in bargaining games', *Games and Economic Behavior*, vol. 7, pp. 341–61.

Hoffman, E., McCabe, K. and Smith, V. (1996). 'Social distance and other-regarding behavior in dictator games', *American Economic Review*, vol. 86, pp. 653–60.

Knack, S. and Keefer, P. (1997). 'Does social capital have an economic payoff? A cross-country investigation', *Quarterly Journal of Economics*, vol. 112, pp. 1251–88.

LaPorta, R., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R. (1997). 'Trust in large organizations', *American Economic Review: Papers and Proceedings*, vol. 87, pp. 333–8.

Levine, D. K. (1998). 'Modeling altruism and spitefulness in experiments', *Review of Economic Dynamics*, vol. 1, pp. 593–622.

Rabin, M. (1993). 'Incorporating fairness into game theory and economics', *American Economic Review*, vol. 83, no. 5, pp. 1281–302.

Roth, A. E. (1995). 'Bargaining experiments', in (J. Kagel and A. E. Roth, eds.), *Handbook of Experimental Economics*, Princeton: Princeton University Press, pp. 53–348.

Shapiro, D. L., Sheppard, B. H. and Cheraskin, L. (1992). 'Business on a handshake', *Negotiation Journal*, vol. 8, pp. 365–77.

Sheppard, B. H. and Tuchinsky, T. (1996). 'Micro-OB and the network organization', in (R. M. Kramer and T. R. Tyler, eds.), *Trust in Organizations. Frontiers of Theory and Research*, Thousand Oaks: Sage.

Weber, M. (1968 [1922]). *Economy and Society*, (G. Ross and C. Wittich, eds.), Berkeley, CA: University of California Press.

White, H. (1980). 'A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity', *Econometrica*, vol. 48, pp. 49–72.

Yamagishi, T. and Yamagishi, M. (1994). 'Trust and commitment in the United States and Japan', *Motivation and Emotion*, vol. 18, pp. 129–66.