Behavioural Development Economics: Lessons from Field Labs in the Developing World

JUAN CAMILO CARDENAS* & JEFFREY CARPENTER**

*Facultad de Economía – CEDE, Universidad de los Andes, Bogotá, Colombia, **Department of Economics, Middlebury College, Vermont, USA

ABSTRACT   *Explanations of poverty, growth and development depend on the assumptions made about individual preferences and the willingness to engage in strategic behaviour. Economic experiments, especially those conducted in the field, have begun to paint a picture of economic agents in developing communities that is at variance with the traditional portrait. We review this growing literature with an eye towards preference-related experiments conducted in the field. We also offer lessons on what development economists might learn from experiments. We conclude by sharing our thoughts on how to conduct experiments in the field and then offer a few ideas for future research.*

I. Introduction

Experimental economics offers methods to test many behavioural hypotheses at the core of why, after decades of attempts to induce development with interventions through markets, the state and self-governance, a few countries have escaped poverty while others remain severely poor. These issues now influence the mainstream of development economics (Ray, 1998; Bardhan and Udry, 1999; Hoff and Stiglitz, 2001) and behavioural experiments have begun to demonstrate the importance of testing (and perhaps, replacing) the standard assumptions made about decision-makers. Inspired by our own experiences and those of a growing number of researchers, we highlight some of the lessons that can be learned from conducting behavioural experiments in developing communities. We focus on a topic that has always been at the heart of development economics – individual preferences. We look for insights in four categories of experiments: the propensity to cooperate in social dilemmas; trust and reciprocity; norms of fairness and altruism; and, risk and time preferences.

(1)

Without state support, many developing communities rely on local norms and rules of conduct to provide public goods and regulate the use of common pool resources. Both of these situations can be categorised as social dilemmas in which individual incentives are at odds with group incentives. For example, egoistic individuals could take advantage of the contributions to public goods made by others (that is, 'free ride') but, the community – as a whole – does better when everyone contributes. The obvious question is whether the individuals making these decisions are egoistic? In addition, the more subtle question is whether egoism is acquired – in other words, do egoism and institutions co-evolve? Our review uncovers a considerable amount of cooperation among people in developing communities; however, there is variation and although egoists are not the dominant type, they do exist. One of the most important lessons from this research might be that cooperative predispositions are heterogeneous and theory needs to account for this fact.

Much of the recent interest in social capital is motivated by the ability of norms of trust and reciprocity to substitute for formal institutions and complete otherwise incomplete contracts. Communities rich in trust and reciprocity are thought to be more productive because the stock of these norms allows people to engage in relationships that would not be profitable if people were not reciprocally-minded. An initial response to this hypothesis was to correlate the performance of countries with survey measures of trust and reciprocity (for example Knack and Keefer, 1997). Experiments are now available that allow one to measure these norms more precisely. Our survey of the results of such experiments suggests that the variation in experimental measures of trust and reciprocity are indeed correlated with important economic indicators like the growth rate of GDP, the fraction of the population in poverty, the rate of unemployment and the Gini coefficient.

Developing communities must often adjudicate disputes locally. An interesting question is whether the norms of fairness and altruism that evolve to fill this void vary in some systematic way with, for example, the dominant production technology (for example, individual versus team production) and whether this variation correlates more generally with the economic performance of the communities. We review the distribution experiments conducted in developing communities and, again, find substantial variation in the norms that evolve and the extent to which these norms appear to be stable. In many cases the allocations that tend to be punished as unfair are precisely those allocations that people on the other side of the interaction tend to avoid.

The remaining experiments that we consider, have been constructed to elicit risk and time preferences. This sort of experiment has a long tradition in development literature (starting with Binswanger, 1980) and has been largely motivated by the proposition that impatience and risk aversion might explain why poor people remain poor. Our reading of the experimental literature does not support this proposition. We find little evidence that poor people in developing countries are more risk averse than people in the developed world; however, the impatience results are mixed.

Section two is the heart of our review, where we describe the preference experiments that are now common in the field; and we highlight the important results that might inform the direction of new research on economic development. We are more critical in section three where we point to a few methodological problems with

the current literature. In section three we also offer advice on how to run experiments in the field. We conclude our review in section four with ideas on how to push the frontier of preference research in developing communities.

II. Preference Experiments Conducted in Developing Communities

In this section we describe the standard experimental protocols used to elicit preferences, any universal patterns that we see in the data and, most importantly, lessons of relevance to development economists. Tables 1–5 contain summary statistics from many of the experiments.

The Propensity to Cooperate in Social Dilemmas

Research on the cooperativeness of individuals in developing countries has used three different experiments: the prisoner's dilemma, the voluntary contribution mechanism, and the common pool resource game. Each game sets up a social dilemma for the participants in which one strategy leads to the social optimum, while the dominant strategy (or best response function) leads to a socially inefficient outcome. The *prisoner's dilemma* (PD) is typically conducted as a symmetric two-person game with two strategies: cooperate and defect, where defect strictly dominates cooperate. PD players are often presented with a normal form matrix in which the payoffs are measured in the local currency and, are asked to choose simultaneously whether to cooperate or defect. However, because interpreting matrices can pose difficulties for respondents, the game has also been played in the form of a vignette. The *voluntary contribution mechanism* (VCM) allows players to contribute to a public good, which has the incentive structure of a *n*-person prisoners dilemma. This game is repeated for a number of rounds. At the beginning of each round, participants are given an endowment of tokens that they can place in two accounts: a private account that only benefits the decision-maker and a public account that benefits everyone in the group. The amount contributed to the public account is a measure of the cooperativeness of the participant. Lastly, in the *common pool resource* game (CPR), players cooperate by not extracting too much from a resource that is accessible to all players but rival or subtract (that is, one player's extraction reduces the resources available for others to extract). Since this game usually has the flavour of a non-linear public bad, explaining the pay off function ② does not usually help participants. Instead, they are presented with a table which relates their extraction choice to the aggregate choice of the rest of the group. Participants use the table to identify their best response to any extraction taken by the rest of the group.

Only a minority free ride as a first impulse (Table 1). Approximately one-third of players cooperate in the PD, contributions of half the endowment are common in the VCM and extracting only three-quarters of the Nash level appears to be the norm in the CPR game. However, there is considerable variation in play. College-aged participants in the United States show only moderate rates of cooperation that tend to decline in repeated versions of the VCM, while cooperation rates are higher and sustained among poor participants in Africa and Southeast Asia, suggesting an inverse relationship between norms of cooperation and development; however, the

Table 1. Cooperation in developing countries

Game	Study	Location	Students	Mean cooperation
PD	Cooper et al. (1996)	United States	Yes	22% cooperate
PD	Hemesath and Pomponio (1998)	United States	Yes	25% cooperate
PD	Tyson et al. (1988)	China	Yes	54% cooperate
PD		South Africa	Yes	45% cooperate w/black other
			Yes	37% cooperate w/white other
VCM	Andreoni (1995)	United States	Yes	33% of endowment
VCM	List (2004)	United States	No	32% of endowment – young
			No	43% of endowment – old
VCM	Barr (2001)	Zimbabwe	No	48% of endowment, 52%[a]
VCM	Barr and Kinsey (2002)	Zimbabwe	No	53% of endowment – women
		Zimbabwe	No	48% of endowment – men
VCM	Carpenter et al. (2004a)	Vietnam	No	72% of endowment, 76%[a]
		Thailand	No	61% of endowment, 73%[a]
VCM	Ensminger (2000)	Kenya	No	58% of endowment
VCM	Gaechter et al. (2004)	Russia	Yes	44% of endowment
		Russia	No	52% of endowment
VCM	Henrich and Smith (2004)	Peru	No	23% of endowment
		Chile-Mapuche	No	33% of endowment
		Chile-Huinca	No	58% of endowment
VCM	Karlan (2005)	Peru	No	81% of endowment[b]
CPR	Cardenas and Carpenter (2004)	United States	Yes	79% of Nash extraction
		Colombia	Yes	74% of Nash extraction
CPR	Cardenas et al. (2000)	Colombia	No	72% of Nash extraction
CPR	Cardenas, et al. (2002)	Colombia	No	68% of Nash extraction, 49%[c]
CPR	Cardenas (2003a)	Colombia	No	74% of Nash extraction, 62%[c]
CPR	Velez et al. (2006)	Colombia	No	80% of Nash extraction

Notes: a. without social sanctions, with social sanctions; b. this result is from a threshold public goods game; c. without communication, with communication.

Table 2. Trust in developing countries

Study	Location	Students	Fraction sent	Fraction returned	Return ratio
Berg et al. (1995)	United States	Yes	0.52	0.30	0.90
Burks et al. (2003)	United States	Yes	0.65	0.40	1.31
Ashraf et al. (2005a)	United States	Yes	0.41	0.23	0.58
	Russia	Yes	0.49	0.29	0.80
	South Africa	Yes	0.43	0.27	0.73
Barr (2003a)	Zimbabwe	No	0.43	0.43	1.28
Buchan et al. (2003)	United States	Yes	0.65	0.45[a]	1.35
	China	Yes	0.73	0.50[a]	1.51
	Japan	Yes	0.68	0.50[a]	1.51
	South Korea	Yes	0.64	0.49[a]	1.47
Burns (2004b)	South Africa	Yes	0.33	0.23	0.70
Cardenas (2003b)	Colombia	Yes	0.50	0.41	1.22
Carter and Castillo (2002)	South Africa	No	0.53	0.38	1.14
Castillo and Carter (2003)	Honduras	No	0.49	0.42	1.26
Holm and Danielson (2005)	Tanzania	Yes	0.53	0.37	1.17
	Sweden	Yes	0.51	0.35	1.05
Danielson and Holm (2003)	Tanzania	No	0.56	0.46	1.40
Ensminger (2000)	Kenya	No	0.44	0.18	0.54
Fehr and List (2004)	Costa Rica	Yes	0.40	0.32	0.96
	Costa Rica	No	0.59	0.44	1.32
Greig and Bohnet (2005)	Kenya	No	0.30	0.41	0.82
Johansson-Stenman et al. (2004)	Bangladesh	No	0.46	0.46	1.38
Karlan (2005)	Peru	No	0.46	0.43	1.12
Koford (2001)	Bulgaria	Yes	0.63	0.46	1.34
Lazzarini, et al. (2004)	Brazil	Yes	0.56	0.34	0.80
Mosley and Verschoor (2003)	Uganda	No	0.49	0.33	0.99
Schechter (2004)	Paraguay	No	0.47	0.44	1.31
Wilson and Bahry (2002)	Russia	No	0.51	0.38	1.15

Notes: a. this figure differs from Buchan et al. (2003) because they include the second-mover's endowment in the amount of money available to send back.

Table 3. Fairness in developing countries (ultimatum game studies)

Study	Location	Students	Mean proposal	Rejection rate
Carpenter et al. (2005a)	United States	Yes	0.41	0.05
		No	0.45	0.07
Cameron (1999)	Indonesia	Yes	0.42	0.10
Gowdy et al. (2003)	Nigeria	No	0.43	0.01
Henrich et al. (2001)	Peru – Machiguenga	No	0.26	0.05
	Tanzania – Hadza	No	0.40, 0.27	0.19, 0.28
	Bolivia – Tsimane	No	0.37	0.00
	Ecuador – Quichua	No	0.27	0.15
	Mongolia – Torguud	No	0.35, 0.36	0.05[a]
	Chile – Mapuche	No	0.34	0.07
	PNG – Au	No	0.43, 0.38	0.27, 0.40
	Tanzania – Sangu	No	0.41, 0.42	0.25, 0.05
	Zimbabwe	No	0.41, 0.45	0.10, 0.07
	Ecuador – Achuar	No	0.42	0.00
	Kenya – Orma	No	0.44	0.04
	Paraguay – Ache	No	0.51	0.00
	Indonesia - Lamelara	No	0.58	0.00
Henrich et al. (2006)	United States	Yes	0.41	0.42[b]
	United States	No	0.48	0.71[b]
	Kenya - Maragoli	No	0.25	0.96[b]
	Kenya – Samburu	No	0.35	0.10[b]
	Kenya – Gusii	No	0.40	–
	Ghana – Accra City	No	0.44	0.33[b]
	Tanzania – Hadza	No	0.26	0.42[b]
	Tanzania – Isanga	No	0.38	0.10[b]
	Siberia – Dolgan	No	0.43	0.35[b]
	PNG - Au	No	0.44	0.43[b]
	PNG – Sursurunga	No	0.51	0.69[b]
	Fiji – Yasawa	No	0.40	0.15[b]
	Bolivia – Tsimane	No	0.27	0.03[b]
	Colombia – Sanquianga	No	0.48	0.30[b]
	Ecuador – Shuar	No	0.37	0.10[b]

Notes: PNG = Papua New Guinea. Two entries in a cell indicate two different samples in the same population.
a. second rejection rate not reported in the original; b. strategy method used so we report the probability that the lowest positive offer (10%) would be rejected.

pattern is not perfect. For example, cooperation among poor slash and burn horticulturalists in Peru is quite low. At the same time, these people and their, relatively, uncooperative neighbours in Chile, do live extremely remotely which suggests that, controlling for the frequency of interactions with strangers (see Henrich, 2000), there may be some relationship between development and the evolution of cooperative norms.[1]

What other regularities are there? It appears that students are less cooperative than other participants. However, the difference may be age driven and not really by schooling. Looking only at non-students, List (2004) finds that older people are more cooperative in the United States, as do Gaechter et al. (2004) in Russia and Carpenter et al. (2004b) in southeast Asia. Allowing participants to socially sanction

Table 4. Altruism in developing countries (dictator game studies)

Study	Location	Students	Mean allocation
Carpenter et al. (2005a)	United States	Yes	0.25
		No	0.45
Ashraf et al. (2005a)	United States	Yes	0.24
	Russia	Yes	0.26
	South Africa	Yes	0.25
Burns (2004a)	South Africa	Yes	0.26
Cardenas and Carpenter (2004)	United States	Yes	0.27
	Colombia	Yes	0.19
Carter and Castillo (2002)	South Africa	No	0.42
Castillo and Carter (2003)	Honduras	No	0.42
Holm and Danielson (2005)	Tanzania	Yes	0.24
	Sweden	Yes	0.28
Ensminger (2000)	Kenya	No	0.31
Gowdy et al. (2003)	Nigeria	No	0.42
Henrich et al. (2006)	United States	Yes	0.32
	United States	No	0.47
	Kenya – Maragoli	No	0.35
	Kenya – Samburu	No	0.40
	Kenya – Gusii	No	0.33
	Ghana – Accra City	No	0.42
	Tanzania – Hadza	No	0.26
	Tanzania – Isanga	No	0.36
	Siberia – Dolgan	No	0.37
	PNG - Au	No	0.41
	PNG – Sursurunga	No	0.41
	Fiji – Yasawa	No	0.35
	Bolivia – Tsimane	No	0.26
	Colombia – Sanquianga	No	0.44
	Ecuador – Shuar	No	0.35

Note: PNG = Papua New Guinea.

free riders also appears to have a noticeable effect on contributions. There is some variation in the sanctioning technology but in all three cases (Barr, 2001; Carpenter, et al. 2004a), cooperation increases when social sanctions are allowed.[2] In the CPR game, simply allowing participants to discuss the game between rounds has an effect similar to social sanctions. For example, Cardenas et al. (2002) show that discussion between rounds of the experiment can reduce extraction dramatically, the effect is lasting and simple discussion leads to conservation levels that are better than when imperfect external regulation is imposed. Carpenter et al. (2004b) regress contribution levels in a VCM on measures of how often participants have informal conversations with their neighbours and find that more chatty Vietnamese participants were also more cooperative.

The Ashraf (2005) experiment complements the monitoring and punishment studies. In this experiment, married couples of Filipinos were asked to make family saving decisions. People were allocated money and have to decide whether to deposit it in an account that only the decision-maker can benefit from, or in one that benefits the entire family. Disclosure of decision matters were only for the author. However,

Table 5. Risk and time preferences in developing countries

Decision task	Study	Location	Students	Mean behaviour
Risk: Accept/reject lotteries	Holt and Laury (2002)[b]	USA	Yes	$0.68 <$ CRRA[a] < 0.97
Risk: Choose lottery	Binswanger (1980)[b]	India	No	CRRA[a] $= 0.71$[c]
Risk: Choose lottery	Wik and Holden (1998)[b]	Zambia	No	$0.81 <$ CRRA[a] < 2.0
Risk: Choose lottery	Barr (2003b)[b]	Zimbabwe	No	$0.32 <$ CRRA[a] < 0.81
Risk: Choose lottery	Harrison et al. (2005)	Ethiopia	No	CRRA[a] $= 0.05$[c]
		India		CRRA[a] $= 0.84$[c]
		Uganda		CRRA[a] $= 0.17$[c]
Risk: Certainty equivalent	Barr and Packard (2000)	Chile	No	CE[d] $= 0.57$
Risk: Certainty equivalent	Henrich and McElreath (2002)	Chile and Tanzania	No	CE[d](Mapuche) $= 0.7$
				CE[d](Huinca) $= 0.4$
				CE[d](Sangu) $= 0.68$
Risk: Accept/reject lotteries	Jimenez (2003)	Spain	Yes	$0.40 <$ CRRA[a] < 1.25
Risk: Choose lottery	Nielsen (2001)	Madagascar	No	CRRA[a] $= 0.32$
Risk: Bet on a die roll	Schechter (2005)	Paraguay	No	CRRA[a] $= 2.57$
Time: Accept/reject delays	Coller and Williams (1999)	USA	Yes	$17\% <$ MIDR[f] $< 20\%$
Time: Accept/reject delays	Harrison et al. (2002)	Denmark	No	IDR[e] $= 28\%$[c]
Time: Questionnaire	Barr and Packard (2000)	Chile	No	IDR[e] $= 43\%$
Time: Accept/reject delays	Kirby et al. (2002)	Bolivia	No	MHIDR[g] $= 12\%$
Time: Choose delay	Nielsen (2001)	Madagascar	No	IDR[e] $= 117\%$
Time: Accept/reject delays	Pender (1996)	India	No	MIDR $> 50\%$

Notes: a. CRRA is the measure of constant relative risk aversion; b. high stakes; c. controlling for demographics; d. CE is the mean certainty equivalent constructed from individual fixed effects in regressions with eight observations per participant that varied the odds of the high and low outcomes. This measure is expressed as a fraction of the high payout; e. IDR is the estimated individual discount rate; f. Coller and Williams (1999) report the median individual discount rate (MIDR) because their analysis is sensitive to cut-offs and the distribution of responses in right-skewed; g. Kirby et al. (2002) report median hyperbolic individual discount rates based on the function $PV = A/(1 + kD)$ where A is the reward, k is the hyperbolic discount rate, and D is the delay in days.

when their wives can find out what they have done, men allocate as much money to the family as women do but, when they can hide their choices, men are likely to keep more. These results suggest that models of household bargaining are incomplete and provide a strong lesson about the use of targeting policy in the Philippines. Cash transfers and grants, intended to benefit children, for example, should either be given directly to mothers or mothers should know when the family has been given money.

A lesson that is starting to emerge from the collection of these cooperation experiments, is that there seems to be a relationship between the existence of formal institutions and the form of norms and social preferences that dictate behaviour. Societies in which formal institutions are missing or weak, develop and rely on pro-social norms and preferences as behavioural benchmarks more than other societies in which institutions are strong. This obviously makes sense: I do not need to feel very cooperative towards you if we can write an enforceable contract but when we cannot, then you demonstrating your cooperativeness matters a lot. If informal enforcement mechanisms are provided locally and cooperation is enforced by social ostracism or mutual monitoring, cooperation may be sustainable.

At this stage, the few observations of the interaction between institutions and cooperation suggest more questions than they resolve. For example, while Cardenas et al. (2000) and Carpenter and Seki (2005) provide some evidence that formal institutions can crowd out social preferences, the dynamics and, in particular, the process of coevolution is not understood very well. Given the possible endogeneity of norms and preferences, there may also be a stability-performance trade-off. There is now plenty of evidence (for example, Ostrom, 1990) suggesting that informal institutions may outperform formal institutions because local solutions are often better informed but, if local norms and rules are less stable than enforced laws this may not hold.

Other, more concrete, lessons can also be learned from cooperation field experiments. First, the cross-national results on the relationship between age and cooperativeness have implications for fostering collective action, volunteering and, perhaps, fund-raising. Second, the strong results on communication and social sanctioning should inform the design of institutions. As the data from Barr (2001) suggest, social shaming can have a big effect on behaviour. If one thinks of microcredit as a form of social dilemma, then the success of group lending is not so surprising given these results.

Group composition is an important predictor of cooperation. For example, Cardenas (2003a) shows how mixing economic classes affects play in a CPR game. Groups composed of mostly poor people conserve common property better than groups which are mixed between poor people and more affluent local property owners. This evidence is contrary to Olson's (1965) privileged group hypothesis. Likewise, Cardenas and Carpenter (2004) show that mixed groups of students from different countries in a CPR game perform noticeably worse than homogenous groups and that these differences are partially explained by conservation attitudes.

At a more basic level, Carpenter et al. (2004a) show how the gender composition of groups affects the level of cooperation among slum dwelling participants in the VCM and this effect appears to depend on location. In Vietnam, homogeneous groups of women are more cooperative but, in Thailand it is the men that are more cooperative. If the effects of asymmetries that we document in experiments are

externally valid, then policies that seem obvious may actually backfire. Extrapolating
from Cardenas (2003a), for example, policies that increase the market value of
labour for those people who extract from real CPRs may actually lead to worse
management. Instead of relying less on extraction from CPR and, therefore,
extracting less, people may just place less value on preserving the resource.

Trust and Reciprocity

In the Berg et al. (1995) investment or *trust game* (TG), two players are endowed
with money as a show-up fee (typically around $10). The first-mover is given the
chance to send as much of her endowment to an anonymous second-mover as she
wishes. The experimenter triples the amount of money sent so that sending money is
socially efficient. The second-mover then sends back as much he wishes.[3]
   The subgame perfect prediction is straightforward. The second-mover has no
incentive to send any money back and therefore, realizing this, the first-mover should
not invest anything in the partnership. Despite this prediction, Table 2 shows that in
Berg et al. (1995), on average, first-movers send 50 per cent of their endowment and
second-movers return 30 per cent of what they receive. Despite deviating from theory
in a pro-social direction, sending money is still a bad investment for first-movers
because they tend to recover only 90 per cent of what they send. In a replication at an
institution with a significantly smaller and more homogenous student population
however, Burks et al. (2003) finds that investing pays off (31% average return).
   Figure 1 and Table 2 summarise our survey of TG behaviour. The first two letters
in Figure 1 indicate the country in which the experiment was conducted and the
name in parentheses indicates the first author of the study. In general, we see that, as
with the cooperation experiments, average play is nowhere near the prediction based
on egoistic preferences. There is an upward sloping relationship between trust and
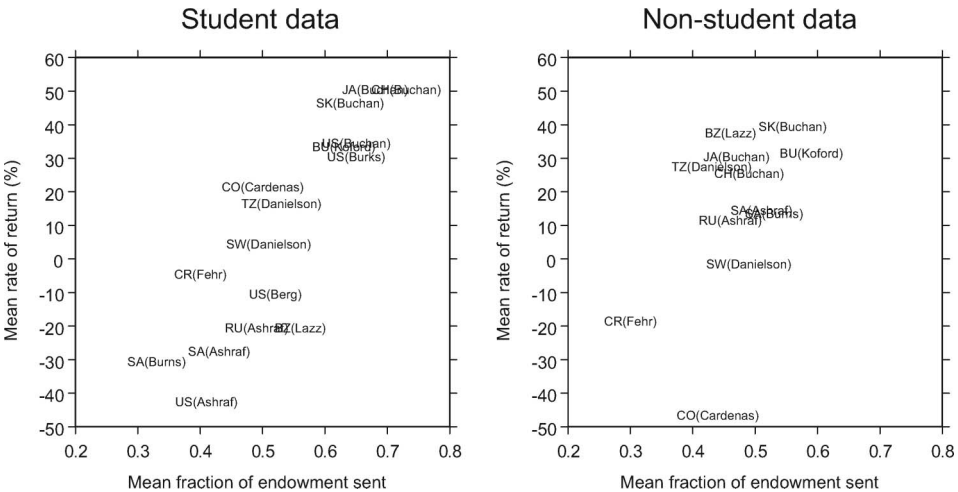reciprocity, suggesting the possibility of multiple trust-trustworthiness equilibria.

Figure 1. An overview of trust game behaviour by country and subject population

At one extreme, the South African students in the Ashraf et al. (2005a) study do not send very much as the first-mover and return significantly less than what is sent to them as second-movers. At nearly the other extreme, Tanzanian non-students in the Danielson and Holm (2003) study tend to send more than half of their endowments and send back a return of 40 per cent, on average. 475

The norms that communities settle on also differ in efficiency. Greig and Bohnet (2005), for example, show that the underlying norm in the slums of Nairobi is *balanced reciprocity*, by which one simply repays an investment as if it were an interest-free loan. By contrast, most of the data from developed countries support the norm of *conditional reciprocity* in which the two parties see the relationship more 480 as a partnership in which both players accrue profit. In the first case there is no relationship between trust and reciprocity, as measured by the experiment but, in the second, the two are positively correlated. What is important for development, however, is the fact that conditional reciprocity is more socially efficient. Other examples of measured community-level differences in the underlying norms driving 485 behaviour can be seen in Danielson and Holm (2003) and Carter and Castillo (2003).

Without the Kenyan outlier, the non-students also seem to demonstrate more trustworthiness. The simple regression of the rate of return on the fraction sent, a student dummy and the interaction of the two shows that the student intercept is significantly lower ($p < 0.05$) and that the student gradient is steeper ($p < 0.10$). This 490 supports what was noted about Russia by Gaechter et al. (2004): student trust experiments should also be seen as lower bounds on pro-social behaviour.

There are also a few methodological lessons to be learned from these trust experiments. First, it seems pretty unambiguous that second-mover behaviour in the TG measures trustworthiness or reciprocity, more generally; however, first-mover 495 behaviour might be harder to interpret. For example, to what extent do first-movers see their transfer as a donation without any expectation of a return? To examine this potential altruism, experimenters have followed Cox (2004) and run dictator games, in addition to TGs. In the *dictator game* (DG), the first-mover simply makes a transfer to a passive second-mover. As there is obviously no self-interested reason to 500 transfer money (the game is anonymous), transfers are interpreted as measures of altruism. Assuming that trust and altruism are additively separable, then pure trust is just the difference between the transfer in the TG and that in the DG. We have identified five experiments that conduct both TGs and DGs. Three of them were conducted in South Africa and find that the amount of 'pure' trust is between 21 per 505 cent (Carter and Castillo, 2003) and 36 per cent (Ashraf et al. 2005a), with Burns (2004a) finding more pure trust (49%) among high school students. Elsewhere in the world the estimates also hover around 50 per cent of the amount sent.[4]

There also now seems to be some consensus that the standard trust question used in the General Social Survey and the World Values Survey (WVS): 'Generally 510 speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?' is a better measure of trustworthiness than trust. Following Glaeser et al. (2000), Lazzarini et al. (2004) and Johansson-Stenman et al. (2004) regress both trust and trustworthiness on the survey trust response; and, finds that the survey question correlates much higher with trustworthiness than with trust. 515 In other words, those people who state higher levels of trust in strangers are actually more trustworthy. One way to look at this result is as a test of the validity of survey

trust questions. A related question asks about the validity of the measures elicited in the TG. The good news is that Karlan (2005), in a Peruvian TG study with participants in a group lending programme, does find correlations between game behaviour and loan repayment behaviour; however, the results are mixed. On one hand, second-mover 'trustworthiness' behaviour appears to be a robust predictor of repaying one's debt but, on the other hand, more 'trusting' first-movers are actually less likely to repay. Like Schechter (2004), Karlan concludes that first-mover behaviour might be confounded by risk preferences. Risk-takers are more prone to taking bad risks that leave them unable to repay their loans and these same people appear more trusting in the TG.

Since the related literature on social capital (see Knack and Keefer, 1997) has focused on trust and trustworthiness, there has been more interest in running the trust game in developing communities. This relative wealth of data means we can go a little further in analysing the links between the behavioural measures and economic outcomes, though such analysis remains tentative.

The first step is to examine the link between the WVS measure of trust used by Knack and Keefer and the experimental measures of trust. As the WVS does not cover six of the countries in the sample, the correlation is based on only 12 observations. However, the results are encouraging. The correlation is positive, $rho = 0.51$ and is significant at the 10 per cent level, indicating that countries with more trust as measured by the WVS, also demonstrate more trust (that is, send more) in the TG.

The second step is to look for correlations between our experimental measures of trust and various economic indicators: GDP growth rate, percentage of the population in poverty, Gini coefficients and unemployment rates (Figure 2).[5] In each case, the correlations are significant at the 5 per cent level or better. Countries with higher growth rates are associated with more trust ($rho = 0.51$, $p = 0.02$); countries with less poverty are associated with more trust ($rho = -0.66$, $p < 0.01$); countries in which the division of economic gains is more unequal are associated with less trust ($rho = -0.48$, $p = 0.04$); and, higher unemployment is associated with less trust ($rho = -0.64$, $p < 0.01$).

Fairness and Altruism

There are two ways to think about distributional norms that may influence dyadic interactions. In the simplest case, norms of altruism dictate how generously one person must treat another when the second person has little or no power to control the outcome. These norms govern many philanthropic acts. Things are a little more complicated, however, when the second person has enough power to retaliate against perceived injustices. To differentiate the norms that dictate behaviour in these situations, we use the term, 'fairness'. Experimenters have developed two simple games to measure norms of fairness and altruism. One game, the *dictator game*, was defined in the previous section. The second game is the *ultimatum game* (UG). In the *ultimatum game*, two players are provisionally allocated a pie to split. The first-mover (proposer) offers a share to the second-mover (responder) who accepts or rejects the offer. Accepted offers are implemented; and rejections result in both players receiving nothing. No responder will choose a rejection threshold
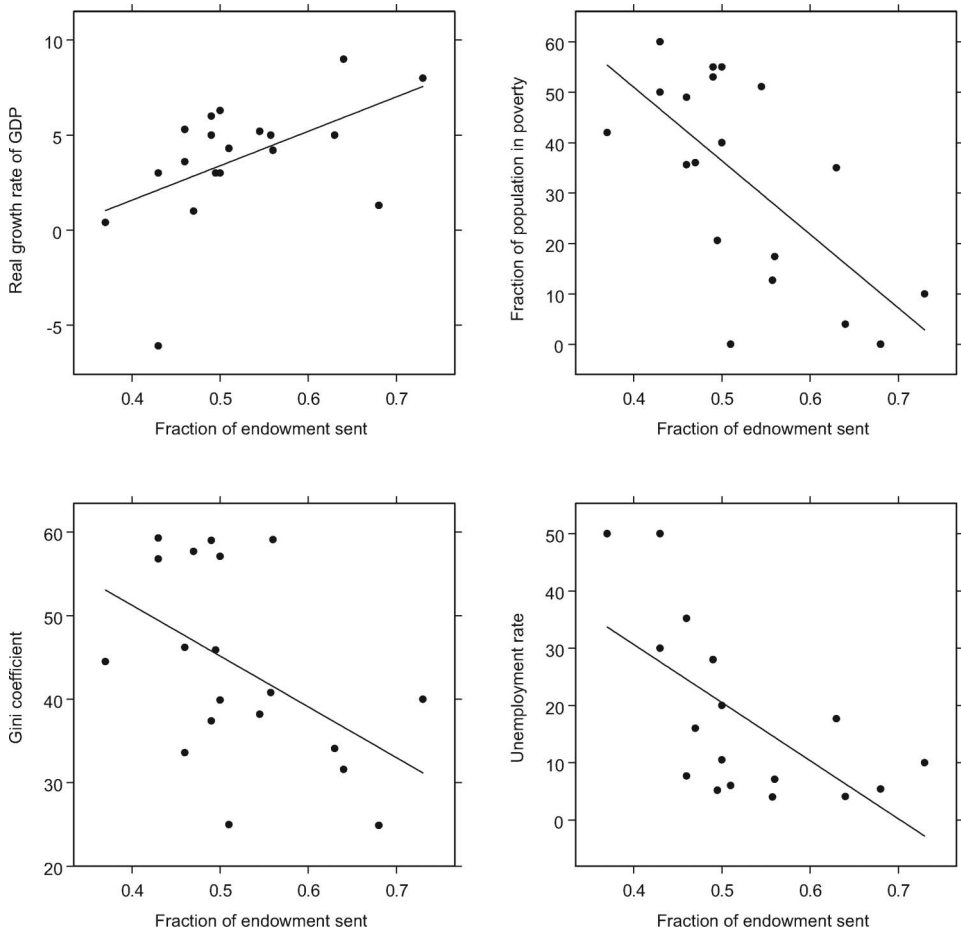
Figure 2. Do experimental measures of trust correlate with economic indicators?

larger than zero because she could do better by accepting lower offers. That is, rejecting should be an empty threat. Knowing this, proposers need not offer more than some small amount. This game has been played hundreds of times in developed countries and, while there is some variation in behaviour across countries (see Roth et al., 1991), most behaviour deviates from the subgame perfect equilibrium (of very small offers) in systematic ways. What is interesting from a development perspective is the variety of distributional norms that arise, whether or not these norms are supported by rejection behaviour and the economic factors that determine the norms.

Fairness behaviour in the UG is summarised in Table 3 and, Table 4 summarises altruism behaviour in the DG. In each case, the mean allocation to the second person is substantially greater than zero and, in the UG, low offers are routinely rejected which suggests that fairness norms are enforced. If you focus on the subset of cases in which both UG and DG data exist, you can see that while students in the US offer

considerably less when the second person cannot punish, this is not true in general. In Carpenter et al. (2005a) and Henrich et al. (2006), students offer slightly more than 41 per cent of the pie in the UG, on average but, only, 25 and 32 per cent in the DG, respectively. However, the difference between the mean UG and DG offers in the developing world tends to be much smaller. This suggests that students in the United States are more sensitive to differences in the strategic environment of a game and the behaviour of other people is more norm-driven. There is another dimension to the Henrich et al. (2006) data that shows the power of local norms. In some cases, participants in the field reject offers that are too high in addition to offers that are too low.

How do norms and institutions interact? There is little research directly on this issue. In their study of 15 small scale societies, Henrich et al. (2001) find that two factors: the local payoffs to cooperation and the degree of market integration explain 68 per cent of the variation in UG offers. Those societies in which team work is essential to production (for example, the Lamelara whale fishermen) have strong sharing norms; while societies composed of small bands of isolated and independent family groups (for example, the Machiguenga) are not particularly generous towards outsiders (nor do they expect the outsiders to be generous). These data also support Hirschman's (1982) theory of civilising markets. People in societies that are more integrated into markets have more experience dealing with strangers and this experience seems to foster more fairness.[6]

Henrich et al. (2006) repeated experiments in a number of the same locations as the 2001 study. Table 3 shows that the second visit to the Hadza (Tanzania), the Tsimane (Bolivia), and the Au (PNG) resulted in data that look very similar to what was gathered in 2001. While it is not the same people playing the game, this result is a mild version of test-retest reliability and should encourage field experimenters that the behaviour measured in these experiments is somewhat robust. There are two other findings worth mentioning. First, they find that community level differences explain much more of the variation in play than individual differences.[7] This supports the idea that distributional norms are local phenomena and, therefore, may vary with local economic conditions. Second, these data confirm that distributive norms are supported by costly punishment. Not only does the rejection behaviour of second-movers in the UG correlate with mean offers by first-movers, using another version of the DG in which an otherwise unaffected bystander can punish the dictator for low offers, the researchers find that the willingness to punish in this variant of the DG correlates strongly with offers in the normal version of the DG. In other words, people are not hypocrites – those who are fairer are more likely to insist on fairness by others.

While there is some variation in UG and DG play across communities, in many places observations with non-student populations pile up on the 50–50 split (for example, Carpenter et al., 2005a). Although this is interesting in the anthropological sense of a human universal, the number of 50–50 splits limits the extent to which the UD and DG can be used to measure fairness norms. To make the games more useful as instruments, participants need to be 'pushed off' the 50–50 split somehow. Recently, Catherine Eckel, Kate Johnson and Duncan Thomas have developed a *comparative* DG to be used in Mexico. The *comparative* DG was constructed in the spirit of Sahlins (1972), who proposed a set of concentric social distance circles

emanating from the individual. Your family comprises the innermost circle, your tribe or some other ingroup comprises the next circle and so forth. You extend and expect different degrees of fairness and reciprocity to and from individuals in different circles. In the *comparative* DG, you are asked to propose an allocation to a family member, a person in your village and a stranger from a different village. Framing the game in this manner affects behaviour as you would expect. These differences are more informative than the standard one-decision DG.

The last two lessons that we find in the fairness data are that deservingness is a strong predictor of altruism and social preference experiments seem to also be good ways to measure social interaction effects (Durlauf, 2002). In Fong (2005), students play a DG in which the recipient will be a local welfare recipient. The amount given depends on whether the potential recipient indicates her willingness to 'pull herself up by her bootstraps.' Those recipients who appear industrious yield significantly more than those who appear lazy. Similar deservingness effects are found in Branas-Garza (2003) who finds that donations to recipients are approximately six times larger when the recipient is identified as poor and in Burns (2004a), in which South African participants use school status and race to proxy for deservingness. Castillo and Carter (2003) implement trust and dictator games along with a survey to investigate what Manski (2000) calls social interaction effects. They show how to exploit the variation in behaviour and socio-economic characteristics at both the individual and group level to identify whether peer effects influence choices.

Time and Risk

A recurring theme in development literature has been that people in underdeveloped countries are poor partially because they have preferences that are inconsistent with growth. They have high discount rates and are risk averse enough, so that it is impossible for them to save and take the risks necessary to begin to accumulate capital. For example, Irving Fisher wrote,

> A small income, other things being equal, tends to produce a high rate of impatience, partly from the thought that provision for the present is necessary both for the present itself and for the future as well, and partly from lack of foresight and self-control. (Fisher, 1930: 73)[8]

In his innovative field study, Binswanger (1980) noted that risk preference differences are important because policy makers can do something about hindrances to the access of capital but, may be able to do less about the risk attitudes of those whom capital would help. In this subsection we consider the evidence on this conjecture.

Risk experiments fall into two classes which differ in the way that participants register their choices. One class is based on the *Accept/Reject Lotteries* experiment. The most important methodological contribution in this class is Holt and Laury (2002). In this experiment, participants are presented with two columns of pair-wise lottery choices and they must accept one lottery per line and reject the other. Initially, the first column dominates the second in terms of expected payoff and variance in the payoffs but, eventually, as the probability of the high outcome in the

second column increases, the expected value of the second column starts to dominate. Those who are more risk averse will choose the first column for longer. Since this experiment forces participants to choose between two discrete options, their preferences can only be estimated on an interval. On average, Holt and Laury find that student participants exhibited levels of constant relative risk aversion between 0.68 and 0.97 when they ratchet up the size of the possible payouts which, in the context of the lotteries offered, is very risk averse behaviour.

The second class of risk experiments is the *Choose Lottery* experiment in which participants are also presented with a series of lotteries but, in this case, they are asked to pick one from a list which controls for the probability of winning a large prize (that is, they are all determined by the toss of a coin) but varies the high and low payouts and, in doing so, the expected payoff. Depending on how risk averse a participant is, he should trade off expected return for less variability. Binswanger (1980) was the first to conduct this sort of risk analysis, doing so with peasant farmers in India. Whilst he does hypothesise that increases in wealth will be associated with lower risk aversion, this result is not borne out in the regression analysis. Despite the differences in the protocols, Binswanger's average estimate of constant relative risk aversion fits within the bounds of the estimate calculated by Holt and Laury. Expanding the comparisons (Holt and Laury, 2002; Binswanger, 1980; Barr; Harrison et al.; Jimenez, 2003; Nielsen, 2001; Wik and Holden, 1998) gives us a better idea of whether there are differences between people in developed countries and those in developing countries. There is some variation in the results (Table 5) but it is not explained by development. In fact, the upper bounds on the Holt and Laury and Jimenez data are larger than the mean values found in the developing world. Overall, there does not appear to be much support for the idea that poor people in developing countries are more risk averse than richer people in developed countries.[9]

Time preference experiments come in many shapes and sizes. Not only is it difficult to compare studies because of differences in their protocols, there are other issues that confound the comparison of time preference data. First, the reported discount rates are very sensitive to how interval choices are interpreted. For example, most studies use exponential discounting but Kirby et al. (2002) used hyperbolic discounting. Further, the number of times that interest is assumed to be compounded per year (obviously) affects the implied discount rate although this assumption is rarely mentioned. Second, many of these experiments are confounded by the credibility of the researcher. In most experiments, participants are paid on the spot but by their very nature, time preference experiments must ask people to wait for their payments. Normally, the researchers are strangers to the participants and, therefore, the experimental data may be biased towards higher discount rates because the participants have two reasons: a preference for the present and not trusting the experimenter; for choosing a payment today versus a promised payment in the future. Third, the delays between payments vary with each study which adds one more factor to control for in any analysis.

As our developed country benchmarks, consider the estimates of individual discount rates (IDRs) gathered by Coller and Williams (1999) and Harrison et al. (2002). These experimenters have learned from the past and conduct their experiments carefully. For example, to control (at least, partially) for the credibility

problem of many experiments, these two papers employ *front end delays*, which simply mean that there is no promise of money today. Instead, people choose, for example, between money tomorrow and more money in a week. If participants think there is some chance that the experimenters will welch on a promise to pay in the future, they should not expect for them to be more likely to welch on a payment one week from today than on one that is due tomorrow. Due to the asymmetry of their data, Coller and Williams (1999) report median IDRs between 17 and 20 per cent using student data while Harrison et al. (2002), using data from a large sample of Danes, report mean overall IDRs that control for many demographic factors of 28 per cent.

The evidence on whether poor people are more impatient is mixed. Along with their risk aversion experiment, Barr and Packard (2000) gather IDRs using a hypothetical questionnaire. Considering the overall variation in estimates, the mean IDR reported in Barr and Packard (43%), is in the ballpark of the developed world estimates. However, in their fuller analysis, Barr and Packard go on to show that income is marginally significantly associated with discount rates ($p < 0.10$): those with larger incomes do appear more patient. Likewise, Kirby et al. (2002), who gather data from the Tsimane horticulturalists of the Bolivian rainforest, show that discount rates are correlated with age (positively), education (negatively) and income (negatively) but, that they are not correlated significantly with wealth. Considering behaviours instead of outcomes, Neilsen (2001) in a study of peasant farmers on Madagascar, finds a mean IDR of 117 per cent[10] and that people who report living in areas in which a lot of deforestation has occurred, have significantly higher discount rates.

Interpretation of the risk and time preference data is a little harder than with the social preference data because there are competing models in each domain. Most people, for instance, now believe that the curvature of one's utility function depends on whether the gamble is formulated in losses or gains. Prospect theory hypothesises that people tend to be risk averse in gains but, relatively, risk seeking in losses (Kahneman and Tversky, 1979). Obviously, because people in developing countries face real risks and development policies often add to this uncertainty, it is important to categorise behavioural predispositions as accurately as possible. Harrison et al. (2005) allow the data to determine the extent to which conventional expected utility theory versus prospect theory predicts behaviour and find that the two models play roughly equal roles in explaining the behaviour of poor participants in India, Ethiopia and Uganda. One implication of prospect theory motivating choices is that the subjective probability weights that people assign to outcomes may be very different to the objective weights they should use.

In their estimates, Harrison et al. (2005) find that when participants are told that an outcome has a 50–50 chance of occurring, they behave as if the chance was as low as 10 per cent. The authors speculate that recent droughts, in the areas in which the experiments were conducted, may have accounted for this general pessimism about uncertain events. The obvious lesson, however, is that with half the people systematically underweighing the likelihood of good outcomes, for instance, it becomes even harder to rally support for change. The next step is to assess the degree to which this probability bias is endogenous and whether simple, first-step, policies would be helpful in altering attitudes towards change.

800   There are also two competing models to describe intertemporal decision-making. Economics and finance rely on simple exponential discounting. However, many recent studies with students in the developed world have found that models that assume hyperbolic discounting, work at least as well (see the review of Frederick et al., 2002). Compared to the exponential standard, hyperbolic discounters
805   appear to have inconsistent preferences: they appear impatient in the short term but extremely patient when considering decisions that will not have to be taken until some distant, future date. In other words (that is, those of Elster, 1989), hyperbolic discounters suffer from a 'weakness of will.' They know what is good for them in the long run but, when the long run becomes the short run they can not overcome their
810   impulses. What are the implications for development policy? It might be the case that people who suffer from weakness of will just need help committing to more financially sound decisions. As Mullainathan (2004) suggests, rotating savings and credit associations (ROSCAs), might serve this purpose. Each member of a ROSCA contributes an amount to a group fund that goes to one of the members at each
815   meeting. Membership requires people to contribute and takes this money out of the member's hands until it is their turn to receive the fund. While this forced savings pays no interest, it does act as a commitment device that is supported by social pressure to not free ride.
        Is there some relationship between poverty, the lack of formal institutions and
820   hyperbolic discounting? In some sense, this is just a new version of Fisher's argument – is the commitment problem greater in developing communities than anywhere else? Unfortunately because most economists still assume exponential discounting, there is not a lot of evidence; however, we have found three studies of relevance. As a non-student benchmark, Harrison et al. (2002) conclude that, in a
825   nationally representative sample of Danes, there is not a lot of evidence to support hyperbolic discounting. There is heterogeneity but most people have relatively constant discount rates over the one to three year horizon. Kirby et al. (2002) estimate hyperbolic discount rates among Amerindian horticulturalists in Bolivia but because the maximum delay in their experiment is less than 6 months, the difference between
830   the implied hyperbolic and exponential rates of time preference are too small to determine which model fits best. By construction, there would be little evidence that the Amerindians suffer from weakness of will. In a very cleanly designed field experiment, Pender (1996) returns to the same villages in India in which Hans Binswanger conducted his seminal risk experiments. For our purposes, the key result
835   is that while he does find discount rates to be higher in a 7-month experiment than in a 12-month experiment, the 12, 19 and 24-month results are indistinguishable. In other words, there is just not enough evidence to make any conclusions about the relationship between development and the prevalence of hyperbolic discounting.
        In sum, we find little evidence of differences in risk and time preferences between
840   people in developing and developed economies. Binswanger (1980) suggests that access to credit might be much more of a problem than different attitudes to risk. This explanation jibes with other behavioural evidence. For example, neither Barr and Packard (2000) nor Jimenez (2003) find that risk preferences predict whether a person chooses to be self-employed. Either the risk experiments are not very
845   externally valid (which would be contrary to Dohmen et al., 2005) or other factors like access to credit dominate these decisions.

III. Running Experiments in the Field

In this section, we reflect on our experiences and on the protocols used by others to formulate a few lessons about running experiments in the field. Other researchers have discussed some of these issues;[11] however, we think that our choice of topics might be more valuable for development economists. We begin by talking about some of the attributes of the participants and researchers and, then, switch to comments on the procedural details of running experiments in the field.

There are good reasons why students are known as a 'convenience' sample of the population. Students are mostly literate, numerate, are somewhat used to thinking abstractly and are accustomed to being told what to do. All these attributes make the running of decision-making experiments go more smoothly. In the field, however, one should never take literacy or numeracy for granted. Furthermore, experiments can often remind non-student participants of the exams that caused them to leave school as soon as possible. It is in the interest of the researcher to make the experiment seem more interesting and straightforward than a test.

Simple instructions that rely more on pictures, diagrams and examples are more successful than those using complex grammar and pay off functions. Literacy is a large problem but it is numeracy that might be something that field experimenters may stumble on and not be able to recover from. In particular, many people have no idea about the basic laws of probability. For example, Herb Gintis once offered a plausible alternative explanation for why the mean offer in the Henrich (2000) UG run, with the Machegeunga in Peru, was so low. The published reason has to do with culture and the low frequency of interacting with strangers. While this reasoning fits well with the other observations from the 15 small scale societies project, a plausible alternative is that nothing is random to people living in Amazonia. Henrich flipped a coin to see if each participant was going to be the first-mover or the second-mover. A coin flip does not mean the same thing to a person who believes that supernatural forces determine the course of events. Such a person may think that she was chosen to be the first-mover for a reason and, with that mindset, demanding most of the pie does not seem unreasonable. Similar reasoning on the part of the person that does not win the toss, may influence one's willingness to reject low offers. The point is, however, that these participant attributes all pose additional problems in the field.

To illustrate the potential importance of numeracy, we can share preliminary results from a study of truck driver trainees in the United States. With 749 of the targeted 1000 observations collected, Stephen Burks, Jeffrey Carpenter, Lorenz Goette and Aldo Rustichini found that a standard test of numeracy correlates highly with trainee behaviour in a variety of experiments. Based on simple correlation coefficients, more numerate people are less impulsive ($p < 0.01$), more cooperative as the first-mover in a sequential PD ($p < 0.02$), more reciprocal as the second-mover in a sequential PD ($p < 0.01$), better at backward induction ($p < 0.01$) and have lower discount rates in the short run ($p < 0.01$).

We also need to worry about the attributes of the people conducting the experiments. Both the protocol and the experimenter have to be credible. Field experimenters are rightly criticised by anthropologists for 'helicoptering' into a site

to conduct an experiment, after spending next to no time with, and knowing next to nothing about, the participants. It is reasonable to fear that this behaviour affects the credibility of the experimenter and may bias the results. A concrete example may better illustrate this point. As mentioned in section two, time preference experiments are often conducted with front-end delays, precisely because the lack of credibility of the experimenter will bias the results towards making people appear more impatient than they really are. Further, if one's assessment of another's credibility is, in turn, determined by economic circumstances, the bias could easily be stronger in developing countries. One obvious way to mitigate any credibility bias is to spend considerable time in the studied communities. This is often not possible but, at a minimum, teaming up with people who have more credibility might be a close substitute.

Sampling and recruitment is an issue for all of behavioural economics but the complications arising from non-random sampling and non-representative recruitment may be more pronounced in the field. The typical lab recruiting protocol is to run an advertisement in the school newspaper or send out a mass email. Not surprisingly, this method is not likely to result in a representative sample (Zelenski et al., 2003). For that matter, assignment to treatment is also usually not very random. If treatment A happens on Monday and treatment B is scheduled for Tuesday, there could easily be some important unobservable that affects behaviour and is correlated with the schedules of the students. These problems are likely to be exacerbated in the field because recruitment is often even more chaotic. Given that the researcher has spent a lot of grant money to get to the site, recruitment often becomes like big game hunting. The researcher sets up her blind near the 'watering hole' and waits for the prey/participants to show up. As a practical matter, this means that recruitment is much more likely to happen by word of mouth and, therefore, peer effects, for example, might add to the sampling problems. When friends or relatives show up for a particular session and the experimental design is highly sensitive to 'social ties' (for example cooperation, trust, public goods, common-pool resource games), one may prefer to assign them to different sessions.

Given that field participants are likely to find the instructions and protocol challenging, field experimenters should do what they can to improve how well the game is understood. Simple things can help increase the quality of the data gathered in the field. Paper and pen experiments, which are often more difficult to design and take longer to run, are actually preferred to computerised experiments in the field because you do not need to worry about computer literacy, on top of the other participant attributes mentioned above. Reading the instructions aloud, using several examples and providing large posters of the decision sheets greatly improves the understanding of the participants. We find that examples work well but it is hard to decide which examples to include because examples may also prime participants on one strategy or another. The jury is still out on framing. On the one hand, people argue that framing is bad because frames may also cue norm-driven behaviour and we are looking for robust behaviours that may not be situation-specific. On the other hand, mild frames may assure that all the participants are playing the game that the experimenter intended. In other words, it may be more important for everyone to be 'on the same page.'

Henrich et al. (2006) have used the *strategy method* in which participants are asked to provide a full strategy of behaviour instead of responding to whatever stimulus they encounter. A second-mover, for example, in the UG provides a full strategy by stating which offers she will accept and which offers she will reject, rather than accepting or rejecting the offer that she is presented with. The benefit is obvious; the experimenter receives much more information. However, the potential problem with this method of eliciting responses is that many people do not naturally think in terms of strategies and, therefore, it is hard for them to think of what they would do in every possible situation, while it is easy to think about what they will do in the situation in which they find themselves.

In some settings, cash payments are less desirable than payment in durables. In developing communities that are not very well integrated into markets, cash is not particularly appealing while tools and other durables are. Hence, to properly incentivise participants, one should be sensitive to their needs. For example, in a study of the rural poor in Andhra Pradesh, Pender (1996) used rice as payment instead of cash. Regardless of the incentive type, paying on average one to two days wage for a half-day session seems to create the necessary salience for participants in the field.

A final concern is *cross talk*; that is when one set of participants talks about the experiment to another set that have not yet participated. This can be a problem because experiments that happen over a sequence of days may implicitly set up a system of overlapping generations so that later choices are biased by earlier behaviour. Cross talk may also introduce selection problems because it could affect who participates. Preventing cross talk is a major challenge. Some people try to recruit large numbers of participants to run concurrent sessions and, other people build in waiting periods, to separate the people who already have participated from those that have not. In either case, experimenters should keep record of the exact day and time of each session in order to keep track of possible cross talk effects.

IV. The Frontiers of Preference Research in LDCs

We conclude by offering a few thoughts on the future of preference-related experimental research in developing countries. We begin by discussing each of the areas covered in section two before ending with a discussion of a few other related ideas.

We still have limited knowledge of the determinants, stability or implications of social preferences. However, the little that we do know suggests that economic outcomes might be sensitive to the distribution of social preferences. Considering the propensity to cooperate in social dilemmas, we have just begun to understand the relationship between formal institutions (laws, in particular) and informal norms of cooperation. We should know more about the co-evolution of formal and informal rules to promote group welfare.

The experiments we have surveyed suggest that information is an important determinant of individual cooperativeness because most people are, to some degree, conditionally cooperative. Information cues the norm of conditional cooperation directly and affects the expectations that allow simultaneous games to be solved

cooperatively. We should conduct research to better understand the information that helps promote and sustain the social preferences that support cooperative behaviour and collective action.

Once the basic trust game has been run in more countries, it will be possible to attempt replication of the work of Knack and Keefer (1997), Zak and Knack (2001) using behaviourally validated data. Focusing at the community level will also allow economists to test the hypothesised links between trust, trustworthiness and the existence of poverty traps. The poverty trap explanation of low growth is slightly different than the standard linear social capital model. Poverty traps are thought to exist because there are multiple stable trust-growth equilibria. If these models better describe reality, then the key is to shock the system out of the basin of attraction of the low trust–low growth equilibrium.

Rosenzweig (1988) concludes that extended families are often substitute for formal institutions in developing countries. Disputes are more likely to be adjudicated through the family than through a formal system of courts. Bonds of common experience and altruism enable families to transcend the informational barriers that tend to hinder the development of formal institutions. If this is indeed the case, then field experiments should be the obvious way to identify the norms of fairness and altruism at the core of these family structures. Once the norms have been identified, it will be possible to examine if variation in norms accounts for important economic indicators like education attainment and income.

The evidence that we have gathered suggests that poor people are not more risk averse (or risk-loving) than rich people and there is mixed evidence on whether poor people are more impatient than rich people. More data would allow us to lay arguments, such as those of Fisher to rest, once and for all. A standardised protocol should be developed based on the latest research from the field and experiments should be conducted in 'sister cities' straddling the development divide that control for many of the differences (for example, distance to a major city and population density) that confound current comparisons.

Mullainathan (2004) lists a number of development puzzles (for example, poverty traps, non-optimal resource use) that might be explained by hyperbolic discounting. However, the data does not currently exist to test whether poor people are more likely to discount the present more than the future. Not only do we not know the extent to which people in developing communities discount hyperbolically, we also do not have much knowledge about the degree to which this preference is endogenous. In what sense might the conditions of poverty cause hyperbolic discounting? What is the impact of education on the propensity to discount hyperbolically? In the spirit of Ashraf et al. (2005b), what policies and contracts can be offered to help poor people commit to savings, given they may discount hyperbolically?

Finally, in many places, we have pointed to the interaction and tension between formal and informal institutions. We consider this to be another area in which the return to research will be high. People traditionally think of informal institutions as existing before formal institutions are established; however, we hope that this is no longer obvious to the reader. We think that there are plenty of cases in which formal and informal institutions coexist; there are cases in which formal institutions crowd out informal institutions; and, there are even cases in which informal institutions fill

the void following the decline of formal institutions. Experiments can help further analysis of the relationship between formal and informal institutions. Is it actually the case (as proposed by Hirschman, 1982), for example, that markets – as formal institutions – simply replace other informal economic relationships?

Our sense is that most development economists are still wary of the use of experiments. While being strong advocates, we understand the reticence – behavioural economists are only beginning to give other practitioners a reason to care. This issue is often linked to the idea of *external validity*: do the behavioural propensities that we capture in experiments correlate with economic activity outside the field lab? Other economists may also be sceptical of the process of recruitment and sampling. A major methodological step will be to work to make our samples more representative and, to reduce the self-selection bias that can arise from voluntary participation. Making sure that the incentives used correspond to the opportunity cost of participating and, at a minimum, collecting demographics from people who decline participation should help.

As the reader can see, there is already a substantial body of experimental work on preferences in developing countries. While some questions have been answered, the major contribution of the current literature is the establishment of a field methodology. We also think that these methods will now set the stage for new, more policy-oriented research. If policies are aimed at inducing changes in behaviour to improve outcomes, experiments can provide detailed behavioural data about the effects of certain incentives, institutions or information on a specific context or group. In the future, these data may also make it feasible to calibrate and tailor policies at the local level.

We conclude by summarising some policy relevant findings. Experimental work on the distribution of resources within a household is extremely important, given that grantors typically assume conditional cash transfers must be put in the hands of mothers, to get resources to children. An example is the Progressa programme in Mexico. Field experiments will test the assumptions at the foundation of this programme. Must grantors target women because they are more altruistic towards their children or, is it because they are more likely to invest in the future of their children?

How might the heterogeneity of distributional norms affect policy? As Elinor Ostrom pointed out, some policies function on the basis of state or other third party intervention and others arise endogenously (Ostrom, 1998). As her work suggests, the right policy often depends on community attributes like the strength and composition of local norms. In this sense, local distributive customs often correlate with the extent to which intervention helps the situation or makes it worse. Moreover, the relationship between the nature of institutions (formal versus informal) and the potency of an intervention, extends to situations in which trust and reciprocity (section two) support more efficient outcomes.

Experiments on time and risk preferences will also continue to influence policy. Measuring these preferences allows policy-makers to first assess the extent to which preferences hinder accumulation. So far, there is little evidence that people remain poor because they are risk averse and the evidence on the relationship with time preferences is mixed. However, once the preferences have been measured more precisely, the results may point to more effective policies. If it is the case that

preferences act as constraints, then commitment mechanisms like ROSCAs, in the case of impatience, should be explored; and for constraints due to attitudes towards risk, farmers should be educated in the actual riskiness of new technologies, rather than to continue to allow the ambiguity of choice to paralyse them (Engle-Warnick et al., 2006). On the other hand, if preferences do not correlate with poverty, it is time to redouble efforts on providing access to credit, for example.

Notes

1. Of course all this speculation is intended to be provocative and needs to be weighed against other explanations for the variation in play. One that is obvious and should be remembered when comparing many of the results we review, is that protocols are not standardised in these experiments and the details of the game do often matter.
2. Barr (2001) allowed participants to directly voice disapproval while participants in Carpenter et al. (2004a) paid a cost to send a message (a picture of an unhappy face) to the rest of the group, when they were unhappy with contributions.
3. This is how Berg et al. (1995) ran the game but other variations have been seen. For example, Glaeser et al. (2000) do not give an endowment to the second-mover and double, rather than triple, the transfer.
4. First-movers may also view the game as more of a gamble than an exchange. The potential risk confound has been discussed in Karlan (2005) and examined in Schechter (2004). Of course, such a correlation is not too surprising but it does suggest that insuring against risk might be as effective as building trust.
5. There are fewer observations in the unemployment graph because this information is not reported for Tanzania or Uganda.
6. Of course the sort of farmers' markets that the Henrich et al. (2001) participants are used to, are vastly different than the thick anonymous markets common in developed economies. It is not obvious that the effect of market integration is globally monotonic, especially controlling for the size of the market.
7. Gowdy et al. (2003) and Bahry and Wilson (2003) draw similar conclusions.
8. More recently, this fable has been discussed in Lawrance (1991), Moseley (2001), Neilsen (2001) and Ogaki and Atkeson (1977).
9. We point out but do not discuss, two other studies that do not fit neatly into our two categories of risk experiments. Barr and Packard (2000) adopt the Schubert et al. (1999) method for measuring the certainty equivalents of Chilean adults to determine if the self-employed are less risk averse. Schechter (2005) had people in rural Paraguay bet on the roll of a die to test their attitudes towards risk. Perhaps because the protocol is sufficiently different, Schechter finds rather high levels of risk aversion (the average CRRA was 2.57).
10. However, there appear to be a number of issues with Neilsen's experiment. One problem is that the lowest IDR interval is between 0 and 20 per cent. Another problem is that the intervals are so wide that little precision can be expected.
11. See Roth et al. (1991), Harrison and List (2004) and Carpenter et al. (2005b).

References

Andreoni, J. (1995) Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments. *Quarterly Journal of Economics*, 110(1), pp. 1–21.

Ashraf, N. (2005) Spousal control and intra-household decision making: an experimental study in the Philippines. Working paper, Harvard Business School.

Ashraf, N., Bohnet, I. and Piankov, N. (2005a) Decomposing trust. *Experimental Economics*, forthcoming.

Ashraf, N., Karlan, D. and Yin, W. (2005b) Tying Odysseus to the mast: evidence from a commitment savings product in the Philippines. *Quarterly Journal of Economics*, forthcoming.

Bahry, D. and Wilson, R. (2003) Confusion or fairness in the field? Rejections in the ultimatum game under the strategy method. Working paper, Department of Political Science, Rice University.

Bardhan, P. and Udry, C. (1999) *Development Microeconomics* (Oxford and New York: Oxford University Press).

Barr, A. (2001) Social dilemmas and shame-based sanctions: experimental results from rural Zimbabwe. Working paper, Centre for the Study of African Economies.

Barr, A. (2003a) Trust and expected trustworthiness: experimental evidence from Zimbabwe. *The Economic Journal*, 113(489), pp. 614–630.

Barr, A. (2003b) Risk pooling, commitment, and information: an experimental test of two fundamental assumptions. Working Paper CSAE WPS/2003–05, Centre for the Study of African Economies, Oxford University.

Barr, A. and Packard, T. (2000) Revealed and concealed preferences in the Chilean pension system: an experimental investigation. Discussion Paper Series, Department of Economics, University of Oxford.

Barr, A. and Kinsey, B. (2002) Do men really have no shame? Working paper, Centre for the Study of African Economies.

Berg, J., Dickaut, J. and McCabe, K. (1995) Trust, reciprocity and social history. *Games and Economic Behavior*, 10, pp. 122–142.

Binswanger, H. (1980) Attitudes toward risk: experimental measurement in rural India. *American Journal of Agricultural Economics*, 62(August), pp. 395–407.

Branas-Garza, P. (2003) Poverty in dictator games: awakening solidarity. Working paper E2003/50, centrA.

Buchan, N., Croson, R. and Johnson, E. (2003) Let's get personal: an international examination of the influence of communication. Culture, and social distance on other regarding preferences. *Journal of Economic Behavior and Organization*, 51.

Burks, S.V., Carpenter, J.P. and Verhoogen, E. (2003) Playing both roles in the trust game. *Journal of Economic Behavior and Organization*, 51(2), pp. 195–216.

Burns, J. (2004a) Inequality aversion and group identity in segmented society: evidence from South Africa. Working paper, Department of Economics, University of Cape Town.

Burns, J. (2004b) Rebuilding trust in post-apartheid South Africa. Working paper, Department of Economics, University of Cape Town.

Cameron, L. (1999) Raising the stakes in the ultimatum game: experimental evidence from Indonesia. *Economic Inquiry*, 37(1), pp. 47–59.

Cardenas, J.C. (2003a) Real wealth and experimental cooperation: evidence from field experiments. *Journal of Development Economics*, 70, pp. 263–289.

Cardenas, J.C. (2003b) En vos cofio: an experimental exploration on the micro-foundations of trust, reciprocity, and social distance in Colombia. Working paper, Faculty of Economics, Universidad de Los Andes.

Cardenas, J.C. and Carpenter, J. (2004) An inter-cultural examination of cooperation in the commons. Working paper, Department of Economics, Middlebury College.

Cardenas, J.C., Stranlund, J. and Willis, C. (2000) Local environmental control and institutional crowding-out. *World Development*, 28(10), pp. 1719–1733.

Cardenas, J.-C., Stranlund, J. and Willis, C. (2002) Economic inequality and burden-sharing in the provision of local environmental quality. *Ecological Economics*, 1–17.

Carpenter, J. and Seki, E. (2005) Do social preferences increase productivity? Field experimental evidence from fishermen in Toyama bay. IZA discussion paper 1697.

1130

1135

1140

1145

1150

1155

1160

1165

1170

1175

Carpenter, J., Daniere, A. and Takahashi, L. (2004a) Social capital and trust in southeast Asian cities. *Urban Studies*, 41(4), pp. 853–874.

Carpenter, J., Daniere, A. and Takahashi, L. (2004b) Cooperation, trust, and social capital in southeast Asian urban slums. *Journal of Economic Behavior and Organization*, 55(4), pp. 533–551.

Carpenter, J., Burks, S. and Verhoogen, E. (2005a) Comparing students to workers: the effects of social framing on behavior in distribution games, in: J. Carpenter, G. Harrison and J. List (eds) *Field Experiments in Economics*, (Greenwich, CT and London: JAI/Elsevier), pp. 261–290.

Carter, M. and Castillo, M. (2002) The economic impacts of altruism, trust and reciprocity: an experimental approach to social capital. AAE Staff Papers, University of Wisconsin-Madison.

Carter, M. and Castillo, M. (2003) An experimental approach to social capital in South Africa. Working paper, Department of Agricultural Economics, University of Wisconsin.

Castillo, M. and Carter, M. (2003) Identifying social effects in economic field experiments. Working paper, Department of Agricultural and Applied Economics, University of Wisconsin.

Coller, M. and Williams, M.B. (1999) Eliciting individual discount rates. *Experimental Economics*, 2(2), pp. 107–127.

Cooper, R., DeJong, D., Forsythe, R. and Ross, T. (1996) Cooperation without reputation: experimental evidence from prisoner's dilemma games. *Games and Economic Behavior*, 12, pp. 187–218.

Cox, J.C. (2004) How to identify trust and reciprocity. *Games and Economic Behavior*, 46, pp. 260–281.

Danielson, A. and Holm, H. (2003) Do you trust your brethren? Eliciting trust attitudes and trust behavior in a Tanzanian congregation, ????.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J. and Wagner, G. (2005) Individual risk attitudes: new evidence from a large, representative, experimentally-validated survey. IZA discussion paper 1730.

Durlauf, S.N. (2002) On the empirics of social capital. *The Economic Journal*, 112(483), pp. 459–479.

Elster, J. (1989) *The Cement of Society: A Study of Social Order*, (Cambridge: Cambridge University Press).

Engle-Warnick, J., Escobal, J. and Laszlo, S. (2006) Risk preference, ambiguity aversion and technology choice: experimental and survey evidence from Peru, ????.

Ensminger, J. (2000) Experimental economics in the bush: why institutions matter, in: C. Menard (ed.) *Institutions, Contracts, and Organizations: Perspectives from New Institutional Economics*, (London: Edward Elgar), pp. 158–171.

Fehr, E. and List, J. (2004) The hidden costs and returns of incentives: trust and trustworthiness among CEOs. *Journal of the European Economic Association*, 2(5), pp. 743–771.

Fisher, I. (1930) *The Theory of Interest*, (New York: Macmillan).

Fong, C. (2005) Empathic responsiveness: evidence from a randomized experiment on giving to welfare recipients. Working paper, Department of Social and Decision Sciences, Carnegie Mellon.

Frederick, S., Loewenstein, G. and O'Donoghue, T. (2002) Time discounting and time preference: a critical review. *Journal of Economic Literature*, 40(2), pp. 351–401.

Gaechter, S., Herrmann, B. and Thoeni, C. (2004) Trust, voluntary cooperation, and socio-economic background: survey and experimental evidence. *Journal of Economic Behavior and Organization*, 55(4), pp. 505–531.

Glaeser, E., Laibson, D., Scheinkman, J. and Soutter, C. (2000) Measuring trust. *The Quarterly Journal of Economics*, 65(3), pp. 811–846.

Gowdy, J., Iorgulescu, R. and Onyeiwu, S. (2003) Fairness and retaliation in a rural Nigerian village. *Journal of Economic Behavior and Organization*, 52, pp. 469–479.

Greig, F. and Bohnet, I. (2005) Is there reciprocity in a reciprocal-exchange economy? Evidence from a slum in Nairobi, Kenya. Working paper, Kennedy School of Government.

Harrison, G. and List, J. (2004) Field experiments. *Journal of Economic Literature*, 42(December), pp. 1009–1055.

Harrison, G., Humphrey, S. and Verschoor, A. (2005) Choice under uncertainty in developing countries. CeDEx Discussion paper no. 2005–18.

Harrison, G.W., Lau, M.I. and Williams, M.B. (2002) Estimating individual discount rates in Denmark: a field experiment, *American Economic Review*, 92(5), pp. 1606–1617.

Hemesath, M. and Pomponio, X. (1998) Cooperation and culture: students from China and the United States in a prisoner's dilemma, *Cross-Cultural Research: The Journal of Comparative Social Science*, 32(2), pp. 171–184.

Henrich, J. (2000) Does culture matter in economic behavior? Ultimatum game bargaining among the Machiguenga indians of the Peruvian Amazon, *American Economic Review*, 90(Sept), pp. 973–979.

Henrich, J. and McElreath, R. (2002) Are peasants risk-averse decision makers? *Current Anthropology*, 43(1), pp. 172–181.

Henrich, J. and Smith, N. (2004) Comparative experimental evidence from Machiguenga, Mapuche, Huinca and American populations shows substantial variation among social groups in bargaining and public goods behavior, in: J. Henrich, R. Boyd, S. Bowles, H. Gintis, E. Fehr and C. Camerer (eds) *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from fifteen Small-scale Societies*, (Oxford: Oxford University Press), pp. 125–167.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. and McElreath, R. (2001) In search of homo economics: behavioral experiments in 15 small-scale societies, *American Economic Review*, 91(2), pp. 73–78.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D. and Ziker, J. (2006) Costly punishment across human societies, *Science*, 312(23 June), pp. 1767–1770.

Hirschman, A. (1982) Rival interpretations of market society: civilizing, destructive, or feeble? *Journal of Economic Literature*, 20(Dec), pp. 1463–1484.

Hoff, K. and Stiglitz, J. (2001) Modern economic theory and development, in: G. Meier and J. Stiglitz (eds) *Frontiers of Development Economics*, (Oxford: Oxford University Press), pp. 389–459.

Holm, H. and Danielson, A. (2005) Trust in the tropics? Experimental evidence from Tanzania, *The Economic Journal*, 115(503), pp. 505–532.

Holt, C. and Laury, S. (2002) Risk aversion and incentive effects, *American Economic Review*, 92(5), pp. 1644–1655.

Jimenez, F. (2003) Does familiar environment affect individual risk attitudes? Olive-oil producer vs. no-producer households. Documento de trabajo Serie Economia E2003/28, centrA.

Johansson-Stenman, O., Mahmud, M. and Martinsson, P. (2004) Trust, trust games and stated trust: evidence from rural Bangladesh. Working paper, Department of Economics, Goteborg University.

Kahneman, D. and Tversky, A. (1979) Prospect theory: an analysis of choice under risk. *Econometrica*, 47, pp. 263–291.

Karlan, D. (2005) Using experimental economics to measure social capital and predict financial decisions, *American Economic Review*, 95(5), pp. 1688–1699.

Kirby, K., Godoy, R., Reyes-Garcia, V., Byron, E., Apaza, L., Leonard, W., Perez, E., Vadez, V. and Wilkie, D. (2002) Correlates of delay-discount rates: evidence from Tsimane amerindians of the Bolivian rain forest, *Journal of Economic Psychology*, 23, pp. 291–316.

Knack, S. and Keefer, P. (1997) Does social capital have an economic payoff? A cross-country investigation, *Quarterly Journal of Economics*, 112, pp. 1251–1288.

Koford, K. (2001) Trust and reciprocity in Bulgaria: a replication of Berg, Dickhaut and McCabe (1995). Working paper, University of Delaware.

Lawrance, E. (1991) Poverty and the rate of time preference: evidence from panel data, *Journal of Political Economy*, 99(1), pp. 54–77.

Lazzarini, S., Madalozzo, R., Artes, R. and de Oliveira Siqueira, J. (2004) Measuring trust: an experiment in Brazil. Ibmec working paper – WPE – 2004 – 1.

List, J. (2004) Young, selfish and male: field evidence of social preferences, *The Economic Journal*, 114(Jan), pp. 121–149.

Manski, C. (2000) Economic analysis of social interactions, *Journal of Economic Perspectives*, 14(3), pp. 115–136.

Moseley, W. (2001) African evidence on the relation of poverty, time preference and the environment, *Ecological Economics*, 38, pp. 317–326.

Mosley, P. and Verschoor, A. (2003) The development of trust and social capital in rural Uganda: an experimental approach. Working paper, University of Sheffield.

Mullainathan, S. (2004) Psychology and development economics. Working paper, Department of Economics, MIT.

Nielsen, U. (2001) Poverty and attitudes towards time and risk - experimental evidence from Madagascar. Working paper, Royal Veterinary and Agricultural University of Denmark.

Ogaki, M. and Atkeson, A. (1977) Rate of time preference, intertemporal elasticity of substitution, and level of wealth, *Review of Economics and Statistics*, 79(4), pp. 564–572.

Olson, M. (1965) *The Logic of Collective Action*, (Cambridge: Harvard University Press).

Ostrom, E. (1990) *Governing the Commons: The Evolution of Institutions for Collective Action*, (Cambridge: Cambridge University Press).

Ostrom, E. (1998) A behavioral approach to the rational choice theory of collective action presidential address. *American Political Science Review*, 92(1), pp. 1–22.

Pender, J.L. (1996) Discount rates and credit markets: theory and evidence from rural India, *Journal of Development Economics*, 50(2), pp. 257–296.

Ray, D. (1998) *Development Economics*, (Princeton: Princeton University Press).

Rosenzweig, M. (1988) Risk, implicit contracts and the family in rural areas of low-income countries. *The Economic Journal*, 98(Dec), pp. 1148–1170.

Roth, A., Prasnikar, V., Okuno-Fujiwara, M. and Zamir, S. (1991) Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh and Tokyo: an experimental study, *American Economic Review*, 81(5) (Dec), pp. 1068–1095.

Sahlins, M. (1972) *Stone Age Economics*, (Chicago: Aldine).

Schechter, L. (2004) Traditional trust measurement and the risk confound: an experiment in rural Paraguay, *Journal of Economic Behavior and Organization*, forthcoming.

Schechter, L. (2005) Risk aversion and expected-utility theory: a calibration exercise. Working paper, Department of Agricultural Economics, University of Wisconsin.

Schubert, R. and et al. (1999) Financial decision-making: are women really more risk-averse? *American Economic Review*, 89(2), pp. 381–385.

Tyson, G.A., Schlachter, A. and Cooper, S. (1988) Game playing strategy as an indicator of racial prejudice among South African students, *Journal of Social Psychology*, 128(4), pp. 473–485.

Velez, M.A., Stranlund, J. and Murphy, J. (2006) What motivates common pool resource users? Experimental evidence from the field. Working paper, Department of Resource Economics, University of Massachusetts.

Wik, M. and Holden, S. (1998) Experimental studies of peasant's attitudes towards risk in northern Zambia. Working Paper D-14, Agricultural University of Norway.

Wilson, R. and Bahry, D. (2002) Intergenerational trust in transitional societies. Working paper, Department of Political Science, Rice University.

Zak, P.J. and Knack, S. (2001) Trust and growth, *Economic Journal*, 111(470), pp. 295–321.

Zelenski, J., Rusting, C. and Larsen, R. (2003) Consistency in the time of experiment participation and personality correlates: a methodological note, *Personality and Individual Differences*, 34(4), pp. 547–558.