



People's Trust  
The Design of a Survey-based Experiment

John Ermisch and Diego Gambetta<sup>1</sup>

ISER Working Paper  
2006-34

---

<sup>1</sup> The authors' names are listed in alphabetical order, not in terms of differential contribution. John Ermisch is Professor of Economics at The Institute for Social and Economic Research, University of Essex, and Diego Gambetta is Professor of Sociology, University of Oxford, and Official Fellow of Nuffield College.

## Acknowledgement:

\* The research was supported by the Economic and Social Research Council through their grant to the Research Centre on Micro-social Change in ISER. We are particularly grateful to Abigail Barr, Heather Laurie and Thomas Siedler for detailed comments on an earlier draft of this paper. We also thank James Banks, Michele Belot, Robert Ford, Michelle Jackson and Tak Wing Chan for helpful suggestions.

Readers wishing to cite this document are asked to use the following form of words:

**Ermisch, John and Gambetta, Diego (July 2006) 'People's Trust: The Design of a Survey-based Experiment, ISER Working Paper 2006-34. Colchester: University of Essex.**

The on-line version of this working paper can be found at <http://www.iser.essex.ac.uk/pubs/workpaps/>

The Institute for Social and Economic Research (ISER) specialises in the production and analysis of longitudinal data. ISER incorporates

- MISOC (the ESRC Research Centre on Micro-social Change), an international centre for research into the lifecourse, and
- ULSC (the ESRC UK Longitudinal Studies Centre), a national resource centre to promote longitudinal surveys and longitudinal research.

The support of both the Economic and Social Research Council (ESRC) and the University of Essex is gratefully acknowledged. The work reported in this paper is part of the scientific programme of the Institute for Social and Economic Research.

Institute for Social and Economic Research, University of Essex, Wivenhoe Park,  
Colchester. Essex CO4 3SQ UK  
Telephone: +44 (0) 1206 872957 Fax: +44 (0) 1206 873151 E-mail: [iser@essex.ac.uk](mailto:iser@essex.ac.uk)  
Website: <http://www.iser.essex.ac.uk>

© July 2006

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form, or by any means, mechanical, photocopying, recording or otherwise, without the prior permission of the Communications Manager, Institute for Social and Economic Research.

## ABSTRACT

In this paper we present the design of a two-stage experiment which aims to measure trusting and trustworthiness in a representative sample of the British population. In the first part we discuss the shortcomings of the most common design of the ‘trust-game’ experiment in eliciting information about clear and cogent notions of trusting and trustworthiness, and in the second part we present an alternative design, which we call the ‘framed binary trust game’. The basic design will be administered to a sample of 200 subjects who were formerly members of the British Household Panel Survey (BHPS). In the third part of the paper, we extend this design to allow the ‘truster’ to purchase some information about the ‘trustee’ so as to make the experiment a better representation of real-life trust decisions. We plan in a second stage to run the extended experiment on a larger sample of about 1000 subjects.



## **Introduction**

In this paper we present the design of a two-stage experiment which aims to measure trusting and trustworthiness in a representative sample of the British population. In the first part we discuss the shortcomings of the most common design of the ‘trust-game’ experiment in eliciting information about clear and cogent notions of trusting and trustworthiness, and in the second part we present an alternative design, which we call the ‘framed binary trust game’. The basic design will be administered to a sample of 200 subjects who were formerly members of the British Household Panel Survey (BHPS).<sup>1</sup> In the third part of the paper, we extend this design to allow the ‘truster’ to purchase some information about the ‘trustee’ so as to make the experiment a better representation of real-life trust decisions. We plan in a second stage to run the extended experiment on a larger sample of about 1000 subjects.

We know of only two previous instances in which trust-game experiments have been carried out with a representative sample of the population and combined with survey questions (Fehr et. al. 2002; Bellemare and Kroeger 2003). The study by Fehr and his colleagues gave us the original inspiration. In a study subsequent to the 2002 paper, Fehr and colleagues have run their experiment in three consecutive waves of the German Socio-economic Panel (2003-2005), using a sample of about 1500 persons from the 12,000 households in the Panel. The results from this study have not yet been published.

What has become known as the trust-game experiment (TGE) (Berg et. al. 1995, Glaeser et al 2000) consists of two players, often called first and second mover (FM

and SM). FM is given a sum of money and then the choice to pocket it or to transfer any amount of it, from 0 to the whole sum, to SM (who also may receive an initial endowment of money), with whom FM has been randomly matched by the experimenter. If FM decides to transfer some money to SM, this amount is increased by the experimenter – usually twice or thrice the sum – and passed on to SM, who is then asked to decide how much of it, from 0 to the whole amount, to return to FM. Both FM and SM know everything about the game and the payoffs of their choices, but neither is told anything about the other, except that he or she is another human participating in the same experiment. If SM sends back more than FM sent to him, then FM gains, else he loses some of or all the amount that he sent to SM.

“Trust” is then calculated either as the proportion of FMs who send some positive amount or as the average amount sent on by the FMs. “Trustworthiness” is correspondingly calculated as the proportion of SMs who send some money back, sometimes with a threshold that counts only those who send enough to at least pay off the sum sent over, or as the average amount sent back relative to amount sent on by FMs. Other measures are also extracted, such as the proportion of FMs that gains some positive amount relative to what they sent or at least does not lose. This measure reveals whether trusting pays off.

Behavioural experiments that use the TGE, of which there are now many examples (see Camerer 2003 for a review), are valuable for a number of reasons. One is that, contrary to the predictions of standard rational choice theory that assumes agents to be strictly self-interested, they show that both a majority of FMs are ready to ‘trust’

---

<sup>1</sup> We will draw the sample from a larger sample of about 1200 respondents who, for technical and funding reasons, were dropped from the BHPS in 2002. They over-represent low income households,

others and a large proportion of SMs are ready to be ‘trustworthy’ – even in a one-shot game. In this regard, trust-game experiments are part of the broader field of studies currently challenging the behavioural foundations of micro-economic theory.

Another reason is that they allow behavioural measures of both trust and trustworthiness. The latter is particularly important, for while there may be some limited value in asking *attitudinal* questions about trust – of the kind ‘do you think most people are trustworthy?’ –, surveys are useless when one aims to measure trustworthiness: words are cheap – people would respond positively when asked whether they are trustworthy, whether they are or not. For instance, Glaeser et al. (2000) report that the mean answer on a 1-6 scale of agreement to the statement ‘I am trustworthy’ was 5.3 in their sample of Harvard undergraduates, and this index is not significantly related to trustworthy behaviour in the standard TGE by these students, as measured by the ratio of amount returned to the FMs to the amount available to them—the ‘return ratio’. Indeed, attitudinal measures of *trust* are better predictors of trustworthiness in their sample.

The difficulty of measuring trustworthiness has been a major obstacle to the progress of trust research generally, which as a result has been unsatisfactorily one-sided. Countless papers refer to trust, and by this they mean the disposition to trust or ‘trustingness’, without a clear idea of the extent to which people’s trustingness is an idiosyncratic disposition or belief or a response to actual trustworthiness in the reference group people have in mind when answering trust related questions. We know for instance that there are people and countries in which there is greater or

---

and so our first stage results will be representative of low to moderate income people in Britain.

lesser readiness to trust but we do not know to what extent this reflects actual trustworthiness.

However, experiments too have their limitations, one of which is that they are administered mostly to students, who are not representative of the general adult population. If we believe in the limited evidence we have from wider populations, a sample of students may either underestimate or overestimate trustworthiness in the general population. For instance, Fehr et. al. (2002) find that the amount that the SM sends back is positively correlated with age – so ‘trustworthiness’ should be lower among students – but also positively correlated with good health, which should on average be higher among students.

Thus, combining an experiment with a representative sample makes it possible to overcome those shortcomings and obtain sound and representative *behavioural* measures of both trusting and trustworthiness. We want to be able to trust the measures of trust. Furthermore, the data on individual attributes collected for the sample of subjects in the BHPS permits an analysis of the individual-level correlates of both measures. Analyses of the individual attributes correlated in particular to trustworthiness are almost completely lacking (with the exceptions of Fehr et. al. 2002; Bellemare and Kroeger 2003). Surveys lack information on trustworthiness and experiments usually lack a sufficient range and variation of information on individual attributes. We will have both, and plan to conduct an exploratory analysis on how a wide variety of attributes affect both measures.

We have however fundamental objections to the TGE literature, including those studies using representative samples. “Sociologists and psychologists usually object that this game doesn’t capture all there is to trust because the two-person one-shot game does not include relationships, social sanctions, communication, and so many other rich features that may support or affect trust. That’s precisely the point—the game requires *pure* trust” (Camerer 2003: 85). We do not make these objections but a more radical one, namely that the standard TGE lacks the basic features of a trust situation *even in a one-shot case*, and, as a result, as we argue below, this blurs the link with clear notions of trust and trustworthiness *and* introduces confounding effects.

## I

### A critical review of the TGE

Here we provide a critique of the standard TGE, whether used with students or a representative sample of the population. Let us begin by considering a problem that is acknowledged in the TGE literature.

#### *Whose money?*

A crucial condition any TGE has to meet in order to fulfil its aims is to ensure that FM decides whether to send over some money to SM on the basis of his self-interest, that is in order to gain more than he would by keeping all the money. On that basis, he should then decide whether he trusts SM to send enough money back to make his transfer worthwhile. If motives other than self-interest intrude in FM’s decision it will be impossible to establish to what extent trust really mattered.<sup>2</sup>

This condition is harder to meet than it may seem at first. In TGE, FMs are credited with (but do not often physically receive) some money by the experimenter with which to ‘play’, and *play* they may in the sense that they may not feel any sense of property on promised money that seems to descend from the sky. People may well treat money they find in the street or win by luck differently from money they earn or feel as theirs. This lack of effort on FM’s part to gain the money he gets from the experimenter may cause him to consider this situation as one in which he can risk this god-given gift more than he would risk any of *his* money; or could make him feel that he should be generous and share his unexpected loot with the SM, particularly if SM has not received an endowment. If either generosity, fairness or higher-than-usual risk proneness are part of FM motivation, the experiment is likely to lead to an overestimation of trusting. Or, to be precise, if FM sends nothing we can be reasonably sure that he did so because he did not trust SM to return enough if anything. But if he does send some or all of his money over to SM we shall not know whether he wanted to gain *and* trusted SM or simply wanted in one way or another to be nice. What seems to be due to self-interest-cum-trust may be due to other motivations.

The latter difficulty is partially met in some instances of TGE in two ways. In one by giving an equivalent initial ‘gift’ to SM so as not to make FM feel that he is gaining while SM is getting nothing if he decides not to send any money over to him. This should prevent fairness considerations from entering into FM’s decision. The other way has consisted in first administering a ‘dictator game’ to FM so as to have a

---

<sup>2</sup> By self-interest we mean in the game not necessarily in general. One may plan to give his gains to charity, the important thing is that he plays the trust game in order to gain not to give his money away to SM.

benchmark of his generosity or fairness dispositions.<sup>3</sup> The ‘let’s risk it’ attitude, however, is not met by either of these solutions. It is therefore important to think of additional or alternative solutions that make the crucial condition safely met and reduce the risk of intrusion of other motivations.

It is often the case that the experiments emphasise that the money is being paid to compensate the participants for spending time doing the experiment. When the experiment has been carried out with a representative sample (e.g. Fehr et al 2002) the FM respondent may have to wait for some weeks to receive his money, even if he decides not to transfer any of it. This may undermine a sense of ownership of the endowment with which he was credited.

The more difficult problems, however, are different, and to elucidate them we need to consider carefully the essential elements that are required to characterise a situation as one involving trust, to which we shall now turn.

#### *The essential elements of trust decisions*

A basic trust situation is best conveyed by an example. Suppose a person extends a loan to another because she expects him to do his best to repay it when it is clear that she would do better to refuse the loan if he will make no effort to repay it, and when it is also clear that his selfish interest is to make no effort. Then, we shall say, she *trusts him to repay the loan*. In general, we say that a person ‘trusts someone to do X’ if she

---

<sup>3</sup> A dictator game is also played by two players: the dictator must decide how to split some sum of money between himself and the receiver (who is totally passive so it is not exactly accurate to call this

acts on the expectation that he will do X when both know that two conditions obtain: if he fails to do X she would have done better to act otherwise, and her acting in the way she does gives him the opportunity to be selfish and not to do X (Bacharach and Gambetta 2001: 150).<sup>4</sup>

Trust thus entails the exposure of the truster under an explicit or implicit promise by the trustee to reward that exposure in an agreed manner (Dasgupta 1988). Common knowledge of the payoffs is an essential component of the basic trust situation, but in itself is not enough. The trustee must be informed or somehow know first that he is being *trusted* to do something and next about what that *something* he is being trusted to do exactly consists in. A real life trustee knows both that he is a trustee and that he faces a binary choice not a continuous one: one cannot be a little less or a little more trustworthy – one either is or is not. For instance, if you lend someone your car or bicycle, the trustee implicitly promises that he will return it undamaged, and he either does so or not.<sup>5</sup>

In most TGE, although players know the payoffs, SM is not told that he is being trusted nor does he make any promise to act in a certain way as a result. SM may infer that he is being trusted by interpreting the payoffs, but he could infer other things too and fail to think altogether of the situation as one of trust. Compatibly with those payoffs he may think that he is the beneficiary of an act of unilateral generosity on the

---

a “game”). It has sometimes been administered to SMs in order to distinguish between their generosity or fairness and their reciprocity in response to the transfer to them from FM.

<sup>4</sup> These two conditions are posited also by James Coleman (1990: 98) as essential to a trust situation.

<sup>5</sup> Even when the trustworthy act is of a generic and open-ended kind – ‘I am helpful to you because I trust that you will be helpful to me should the occasion arise’ – it still rests on a threshold above which there is a continuous of acts that are deemed trustworthy, but below which the truster considers the trustee’s act to be insufficiently generous to be trustworthy. – ‘now that you have the chance you are not helping me as much as helped you’.

part of FM, and, if he *prefers* to think so, so as not to feel an obligation to return any money, there is nothing in the design that deters him from doing so, contrary to what happens in simple one-shot trust situations in real life. There is nothing in other words that frames the situation as one of trust.

This indeterminacy is aggravated by another feature of the standard non-binary TGE: SM can return *any* sum he chooses – nothing tells him how much he is expected to return.<sup>6</sup> Thus, even if he thinks of the experiment as involving trust, it is left up to him to decide what being trustworthy in that game exactly means. While this may be a useful feature for some purposes, it misses an essential element of how a basic trust situation presents itself in real life. Mutual knowledge of the nature of the trustworthy act is *constitutive* of a trust game. For example, one cannot trust a neighbour to feed one's dog any number of pellets the neighbour is free to choose; he can only trust him to feed the dog an adequate amount – say half a tin a day. If the truster does not make that clear, or otherwise knows that it is clear in the mind of the trustee, a trustee-neighbour who does not feed the truster's dog properly may do so because of lack of competence in dog feeding needs rather than because of lack of trustworthiness. We cannot tell. For lack of trustworthiness to be detectable the trustee must know what the right amount roughly is. And the truster must know that the trustee knows. If then the trustee underfeeds the dog, the truster will have reasons to believe he is untrustworthy and that he did not do what he was expected to do because, say, he wanted to pocket some of the money meant for dog food or because of carelessness. Similarly, when a truster lends some money to a trustee, both know the amount that the trustee needs to return to be trustworthy.

---

<sup>6</sup> Exceptions of discrete, or 'binary', trust games are Snijders and Kerens (1998), Bacharach et al (2001), Bohnet and Huck 2004, Bohnet and Zeckhauser (2004) and Bohnet et al 2005.

How likely it is that SM will think of the standard TGE as involving trust is thus difficult to establish: if FM hands over some money to SM and the TGE instructions indicate that SM can return *any* amount he likes, SM is encouraged to see this as an act of open-ended generosity not of self-interest sustained by trust. SM thus feels free either to take advantage of such an act or to respond to it, and if he does respond to it he can feel free to acknowledge it in a number of ways – by returning a small ‘tip’, or just enough to reciprocate the amount the FM passed on, or he may feel such a sense of warmth towards FM’s generosity to feel compelled to send back much more. Zak (forthcoming) presents evidence that SM’s receipt of money from FM causes a release of a hormone called oxytocin, which makes SM ‘feel good’ toward FM and induces a ‘temporary attachment’ by SM to FM. The amount of money sent back by SM as a proportion of the amount received was found to increase with the amount of oxytocin released. This emotional response by SM is consistent with a response to FM doing something ‘nice’ for him as well as to ‘being trusted.’

Be that as it may, none of these responses, all of which are compatible with the TGE standard design, are fully capturing what we would understand as acts of trustworthiness in a situation in which both parties know that trust is at stake. In order to capture *pure* trust almost *any* trust has been expelled from the standard TGE design. To some considerable extent, to call the standard form of TGE a *trust* game is a misnomer.<sup>7</sup>

---

<sup>7</sup> We acknowledge that the game is sometimes called an ‘investment game’ (Berg et al 1995) or a ‘social dilemma game’ (Fehr et. al. 2002) or a ‘trust and reciprocity game’, but the outcomes are near-universally interpreted in terms of ‘trust’ and ‘trustworthiness’. For instance, the sub-title of the study

### *Framing*

The problem with the standard TGE however is not completely solved only by making sure that both players know exactly how much money FM can expect from SM. Most TGE present the situation in anodyne monetary terms, letting subjects free to make what they wish of it. This lack of frame puts subjects in an artificial situation. They know the payoffs and may even suspect that the experiment is about trust, but they cannot match what is going on in the game with any familiar situation. This opens up the possibility that when participating in TGE they employ reasoning that would be alien in real life trust decisions, and above all reasoning we cannot gauge. They may think of the experiment as some sort of game. And people may be fiercely competitive and self-interested when playing games precisely because they are games, but in life they may be paying their debts and respecting their matrimonial vows. This fear of ‘framing’ the experiment simply ends up generating indeterminate stimuli. Trustworthiness, by contrast, does not exist in the void. There is no such thing as *pure* trust. It is always trust in someone to do *something*, e.g. pay their debts, or look after one’s children.

Even if emotions are mainly involved in rewarding trust, the decision to trust involves anticipating what the trustee will do—whether or not he will be trustworthy. Decision-making often takes the form of pattern matching – referring new situations to familiar ones – rather than explicit weighting of costs and benefits each time we need to make a decision (see Camerer et. al. (2005) for a fuller discussion). If we wish to measure trust in real world situations we need to frame the experiment in ways that

---

by Fehr et. al. (2002) is ‘Examining trust and trustworthiness by integrating behavioral experiments with representative surveys.’ Glaeser et al (2000) is entitled ‘Measuring trust

allow people to use the ‘automatic processes’ they use when confronted with trust situations in real life.

There is evidence from neuroscience (Camerer et. al. 2005) that suggests that the brain figures out how to carry out tasks efficiently. When confronted with a new problem it initially draws heavily on diverse regions, often including the pre-frontal cortex, where relatively inefficient ‘controlled processes’ are concentrated. With further experience with the problem, the brain seems to shift processing gradually toward brain regions and specialised systems that can solve problems automatically and efficiently with low effort. Framing the experiment helps access these systems.

#### *Trust-warranting properties*

Our understanding of trust, which is derived from Bacharach and Gambetta (2001), suggests a further and deeper shortcoming. In all types of TGE, including our own, the trustee’s monetary payoffs are such that a rational self-interested trustee should never fulfil trust, and, expecting that, the rational self-interested truster should never trust. Since however some of the times players trust and fulfil trust, in real life as well as in experiments, there must be something that makes some trustees trustworthy, namely capable of resisting the pull of their self-interested or *raw* payoffs and prefer to be trustworthy even in one shot games. The primary trust dilemma of the truster thus consists in deciding whether the trustee will be motivated by raw monetary payoffs or whether what we may call his *all-in* payoffs differ and make him trustworthy. In other words, to have a problem of trust at all requires that the truster can at least conceive that a trustee could be motivated by self-interest and thus be untrustworthy in the one shot game; but for the truster to even think of trusting he has

to conceive of the possibility that the trustee could be motivated by forces capable of overriding his raw payoffs and, when all things are considered by the trustee, yield payoffs compatible with being trustworthy.

What are these trust-warranting forces that can tame self-interest? There are many candidates, some of which are ‘local’ while others are ‘global’. Someone may be solidly trustworthy as a father and yet an opportunist with his colleagues or the Inland Revenue. Love of children may be a trust-warranting property which sustains trustworthiness only in specific relationships. Honesty, by contrast, is a property that warrants trust in a wide range of encounters. Other general trust-warranting properties include fairness, reciprocity and trust-responsiveness, the feeling of an obligation to fulfil trust that is put in one.

In this regard the standard TGE presents a problem. It is designed in such a way as to tap only into reciprocity or fairness as the properties that can lead SM to return enough money to FM, which while important are not the only ones, not even perhaps the main ones. As we said, we cannot be sure that FM decides, if he does, to pass on some money to SM because he expects to gain more than by keeping it rather than because of unilateral generosity. But suppose FM does decide to pass on some of his money to SM for the ‘right’ reasons, namely that he wants to gain something more. Then all we can say is that in the standard TGE if FM does that it is because he *trusts* SM to make the transfer worth his while in the sense of expecting that SM will feel motivated either by

- reciprocity: ‘you did something good for me and I will do something good for you’, or by

- fairness: ‘you did something that made me gain some money and we should split the gain’.

Trustworthiness however does not count on these properties alone. People who are neither fair minded nor reciprocal can still have reasons to be trustworthy and stick to their promises. Vice versa, reciprocity may often occur without involving trust: if I attend a colleague’s seminar and he does not attend mine, I may feel bad about this, but I cannot really think of his not attending as an act of untrustworthiness, unless we agreed in advance to attend each other’s seminars. Or, similarly, I may invite someone to dinner and he may reciprocate, but I will not necessarily think of his reciprocation as an act of trustworthiness. The standard TGE seems to us to capture more these type of situations than a typical trust situation. To invoke trust- or untrustworthiness the existence and the common knowledge of an *expectation* of reciprocation is essential. The same can be said of fairness: one can care little about fairness, and still be trustworthy because he is honest rather than fair-minded and likes to be trustworthy and do what he *knows* is expected of him. A trust experiment should not rule out trust-warranting general dispositions that in real life can make a trustee trustworthy, and honesty is obviously one of them and an important one at that.

## II

### **The framed binary trust game**

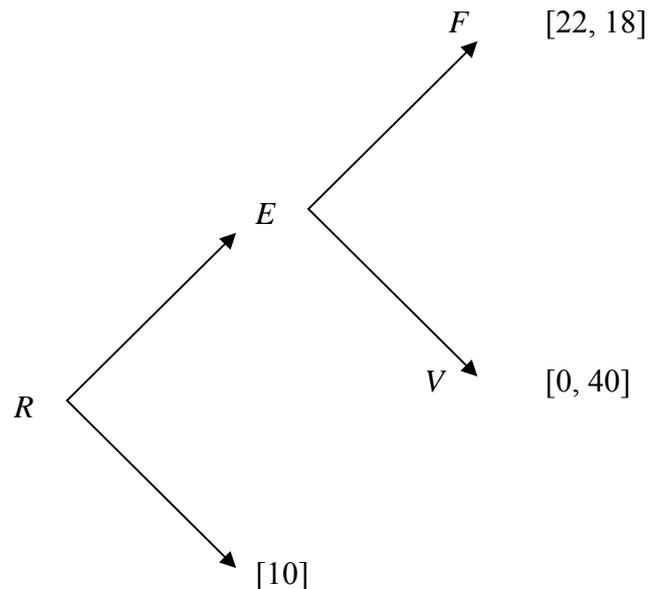
Let us now consider how we propose to tackle the problems we have discussed above with our design.

The game we propose to present to subjects, whom we shall call the truster (R) and the trustee (E), is a basic trust game in binary form. We first describe a treatment in which R must decide whether or not to loan everything he receives, and then briefly describe a second treatment in which he can keep a small part of what he receives.

1. R receives amount  $M$  as compensation for taking part in the interview *cum* experiment. The money, in the form of a post-dated cheque, dated four weeks from the date of the interview (see point 8 below), will be sent to him before the interviewer visits. At the same time R will be notified that, when the interviewer visits, he will have the opportunity to increase his money by ‘*making a loan*’
2. R is offered the binary choice of whether to keep  $M$  or to make a ‘*loan*’ to E of  $M$ . He puts the cheque valuing  $M$  (or a ‘dummy cheque’ provided by the interviewer) in an envelope along with the decision form and seals it; the interviewer does not know his decision.
3. If R does choose to loan,  $M$  is increased by an amount  $J$ .
4. Before making his choice, R is informed that E will be given the choice to either pay back an amount *specified by the experiment* or keep  $M+J$ .
5. Payoffs are such that R would be better off not to make the loan if E chooses to keep it all, but better off to make the loan if E chooses to pay back the specified amount. The problem is that E, in terms of his *raw* payoffs, could do even better by keeping it all; i.e. by not fulfilling the trust placed in him by R.
6. If R chooses not to make the loan the game ends there.
7. If R chooses to make the loan, E is offered a binary choice of whether to pay back the specified amount or keep  $M+J$ .

8. The cheque in point 1 is post-dated in order to make the timing of the payments the same irrespective of whether R chooses to make the loan or not.<sup>8</sup>

The payoffs are illustrated in the following chart (F='fulfill trust', V='violated trust' for M=10, J=30):



We have chosen the amount specified to be paid back in the case that E fulfils trust in a way that makes the payoffs asymmetric. Symmetric payments may encourage fulfilling trust for reasons of fairness, rather than because E does what is expected of her (e.g. see Bacharach et al 2001, Bohnet and Huck 2004 and Bohnet et al 2005).<sup>9</sup> For instance, an equal split may be a focal point. The asymmetric payoffs also discourage E from fulfilling trust because the ‘temptation’ to not fulfil is larger.<sup>10</sup>

In the second treatment, the protocol changes in the following way.

<sup>8</sup> Otherwise Rs with strong time preference (‘impatient’ Rs) would value M+J received 4 weeks hence not much more than M received today, thereby discouraging trusting behaviour. We are grateful to Abigail Barr for pointing out the potential impact of differences in the timing of payments between the two branches of the binary decision.

<sup>9</sup> We are grateful to James Banks for alerting us to this issue.

<sup>10</sup> ‘Temptation’ is usually defined as the payoff gain from violating trust relative to the difference in payoffs between E and R when trust is violated; e.g. see Snijders and Kerens (1998).

1. R receives amount  $M$  as compensation for taking part in the interview *cum* experiment. The money, in the form of two cheques, one for  $M^*$  and the other for  $k$ , will be sent to him before the interviewer visits.
2. R is offered the binary choice of whether to keep  $M$  or to make a 'loan' to E of  $M^* = M - k$ . He puts the cheque valuing  $M^*$  (or a 'dummy cheque' provided by the interviewer) in an envelope along with the decision form and seals it; the interviewer does not know his decision.

The remaining steps are the same as in the previous treatment.

Thus, the two treatments differ in the following way. In the first treatment ( $k=0$ ), R will face a draconian choice of whether to loan all he received or nothing. In the second treatment,  $k$  is a small sum that R can keep even if loans and the initial  $M$  is increased by  $k$ . The rate of return on the loan that R may receive is the same in both treatments. In the tree diagram above, R's payoffs are increased by  $k=2$  in each branch.

Our particular outcome variables are:

- (1) the probability that R makes a loan ('trusts');
- (2) the probability that E pays back the amount specified in the experiment ('fulfils trust').

#### *Key features of our design*

Our design is more realistic than previous games and more precisely captures a clear notion of trust and trustworthiness. This is because:

- (a) R 'gains' his money as compensation for taking part in the interview *cum* experiment and this is sent to him *before* the interviewer's visit. We expect that this should trigger the "endowment effect" and make R treat it more as his own money
-

rather something that he can envisage passing over to E as a gift or because of a win-splitting sense of fairness; this should make R more careful in parting with it.

(b) By forcing R's exposure to a loss of either all or 83% of the sum he receives as a participation payment, we depart dramatically from the standard TGE, in which R can pass on any amount. The distribution of amounts passed on in standard TGE experiments usually covers the entire range, with a distinct modal value of about 50% of R's initial endowment and small 'spikes' at zero and 100% (e.g. see Fehr et. al. 2002; Bellemare and Kroeger 2003 and Barr 2003). In our view, the possibility of transferring any amount allows the intrusion of other motives such as "gift giving", "let's be kind", "let's risk a small part of it", and the distribution of R's transfers is consistent with the operation of these other motives. This confounds these motives with trust, which is a relevant consideration in the decision if and only if R's decision is driven by his self-interested aim to gain more if trust is fulfilled.

(c) The game is framed in terms of a *loan* of his money to E. This is a situation that subjects may easily have encountered in their lives or easily imagine they could encounter, and both R and E know what is expected in a loan situation.

Of course, in real life, the lender and borrower make a contract concerning the terms of the loan, whereas in our design all that the borrower does is to agree to be in the experiment and does not promise to do anything. But we believe that the 'loan frame', though falling short of the contract conditions that regulate trust situations, will still encourage E to put himself in the situation of a person who has agreed to borrow money in the context of an investment opportunity and is expected to behave accordingly.

(d) R knows in advance how much he can expect back from E, and E knows that this is an expectation of R when extending the loan. E thus knows that he cannot be just a

little more or a little less trustworthy; he has to choose whether to be trustworthy or not. Thus, in this binary TGE it is clear what trusting and trustworthiness are, as is common in real life.

A variant of the TGE in Fehr and Rockenbach (2003) uses a solution that could also overcome the problems caused by the freedom of E to return any amount: FM is required to specify a ‘desired back-transfer’ from SM, and SM is informed about it.<sup>11</sup> In this way SM knows also that FM does have expectations and is not simply acting because of generosity. (Oddly, their paper does not report the percentage of SM who make a back-transfer at least as large as the one desired by FM, which would be relevant for trustworthiness if paying such an amount constituted the trustworthy act, suggesting that when they introduced this variant the authors did not have this effect in mind.)<sup>12</sup> However, we chose not to pursue this solution instead leaving the task of specifying the amount expected to us. This has a practical advantage in terms of simplifying the matching between R and E. And, furthermore, has the advantage of standardizing the amount expected in return so that all E respond to the same condition. If one left R free to establish how much he wants to receive from E, too much variation would potentially be introduced, thus making it difficult for us to establish the extent to which trustworthiness was affected by the desired back-transfer established by R – greedy R could be punished by E and receive nothing or not as much as requested back.

---

<sup>11</sup> In a variant, labelled the ‘incentive condition’, the experiment offers the opportunity to levy a fine on SM if he fails to return at least the desired amount. While very important for the purpose of their paper, it is less relevant to the discussion here.

<sup>12</sup> From Table 1 in their paper, it is possible to calculate that among SMs sending a positive amount, the average ratio of the actual back-transfer to the desired one is 0.94, which suggests that a large proportion of SMs who send something (79%) send at least the desired amount.

*The “all-or-nothing” effect*

As noted above, we introduce two payoffs structures, one with  $k=0$  and one with  $k=2$ . Since the rate of return is kept constant in both treatments, this should make no difference in a loan situation, at least in strict rational terms. Yet,

(a) we expect that the treatment in which R has the draconian option ( $k=0$ ) of parting with all his money or keeping it all will *decrease* the chances that R will make the loan. The reason for this prediction is that we expect R to be more inclined to trust when the worse that can happen to them is that they are left with at least something. By contrast when options are draconian, a zero payoff in case trust is not fulfilled will be more discouraging. This may be due to a number of mechanisms, such as risk aversion, fear of loss being triggered more vigorously by the salience of the prospect involving the ‘loss of everything’ etc. Also face saving towards oneself and ‘exploitation aversion’; i.e. R’s fear of being taken advantage of by E could be at work. Recent research suggests that the decision to trust is shaped by exploitation aversion rather than by risk aversion (Bohnet and Zeckhauser, 2004).

At the same time, however,

(b) we expect that the treatment in which R has the draconian option ( $k=0$ ) of parting with all his money or keeping it all, will *increase* the chances that E will pay his due back if R does extend the loan.

The two hypotheses contradict one another for if the latter were true then the former should not be: if R too makes the same hypothesis (b) as we make about E, he should be *more* rather than less inclined to trust. The reason we expect E to be less inclined to

return the money in the non-draconian case is that we think that leaving R with a payoff of 2 will give E a way out, by making him feel less guilty about not fulfilling trust, approximately for the mirror mechanisms that make R more inclined to lend: E will feel that R does not suffer a total loss of face and that he is not leaving him with a worrisome salient *nothing*.

Were (b) to turn out to be true, an interesting consequence would follow for interpreting one result of the standard TGE: it could explain the puzzling finding whereby in many cases the second mover does send only a little back but not enough to cover the sum sent over by the first mover. It would be like giving a guilt-allaying tip.

#### *Risk aversion*

In order to gauge the role of an individual's risk aversion in affecting people's decisions in our experiment, we will ask our participants two simple questions *after they have made their decisions in the experiment*. The first question is the following: 'Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on the scale, where 0 means "unwilling to take risks" and the value 10 means "fully prepared to take risks".' This question was asked in the 2004 wave of the German Socio-Economic Panel (GSOEP). Dohmen et al (2005) conduct a complementary field experiment and find that this measure is a good predictor of actual risk-taking behaviour. They also find that this 'general willingness to take risk' measure does better in predicting risk behaviours than a standard lottery question to measure risk preference.

The second question also comes from the GSOEP. It asks a similar question about willingness to take risks in trusting strangers, again on a 11-point scale. Our own tabulation of the responses on this and the ‘general willingness to take risk’ question indicates that they are moderately correlated ( $r=0.41$ ), but the distributions of responses are different. Overall, people are less willing to take risk in trusting strangers than they are willing to take risk in general. This difference may reflect the exploitation aversion discussed earlier. Responses to these two questions will be related to decisions in our trust experiment.

### *The sample*

The first phase of our study will implement this experiment with a sample of about 200 subjects, who were formerly members of the British Household Panel Survey, but were dropped from the panel for technical and funding reasons. More information about the respondents from which this sample will be drawn is given in Appendix 1. It is not representative of the general population—low-income households and elderly people are over-represented. When we restrict the sample to those age under 65, the age distribution is similar to the entire BHPS, and the sample can probably be taken as representative of households with low to moderate income, for whom we will obtain estimates of our measures of trust and trustworthiness. The focus on lower income people may have its advantages because the sums of money involved in the experiment may be more important to them. The primary purpose of this first phase will be to determine baseline information for the two outcome variables, which will be very helpful in designing the more elaborate (and costly) second phase of our study. If we expect only, say, 50% of Rs to choose to make a loan, we will need to

approach about 135 subjects to be R-players, who will be randomly matched to about 65 E-players.

An advantage of using a sample drawn from people who have participated in the BHPS for a number of years is that they are likely to believe that the field organisation and the managers of the survey can be trusted to carry out the experiment and payments in the way that they promise. Trust responses in a new survey are likely to be contaminated by varying degrees of trust in the organisers of the experiment, in addition to trust in their co-player. While constrained by the small sample, we can also carry out some exploratory analysis of individual correlates of trust and trustworthiness, because we know a lot about the attributes of the sample members.

One person will be selected from each household, and only persons aged under 65 will be sampled. Some basic information, such as current employment and marital status and banded income, will be collected at the time of the experiment. The ‘all or nothing’ treatment will be applied to one-half of the sampled Rs.

### III

#### **FBTG with information**

Our quest to add a more realistic texture to the trust experiment does not end here. In most real life trust situations, some information about the trustee is available or can be obtained at some cost, even in one-shot encounters (see, for instance, Gambetta and Hamill 2005). The notion that playing with anonymous partners helps reveal *pure* trust (Camerer 2003) is unconvincing. What it reveals, is at best R’s belief in the base rate probabilities of trustworthiness of a generic population. And it may be not that

generic either: also in the pure TGE players know that they are playing with a human, that this human belongs to the same language group to which they belong, and that he or she has also accepted to take part in the experiment. Students would guess that this human is almost certainly another student. In some cases the subjects see each other in the waiting room or tell each other that they are participating before they do. Full anonymity is a chimera.

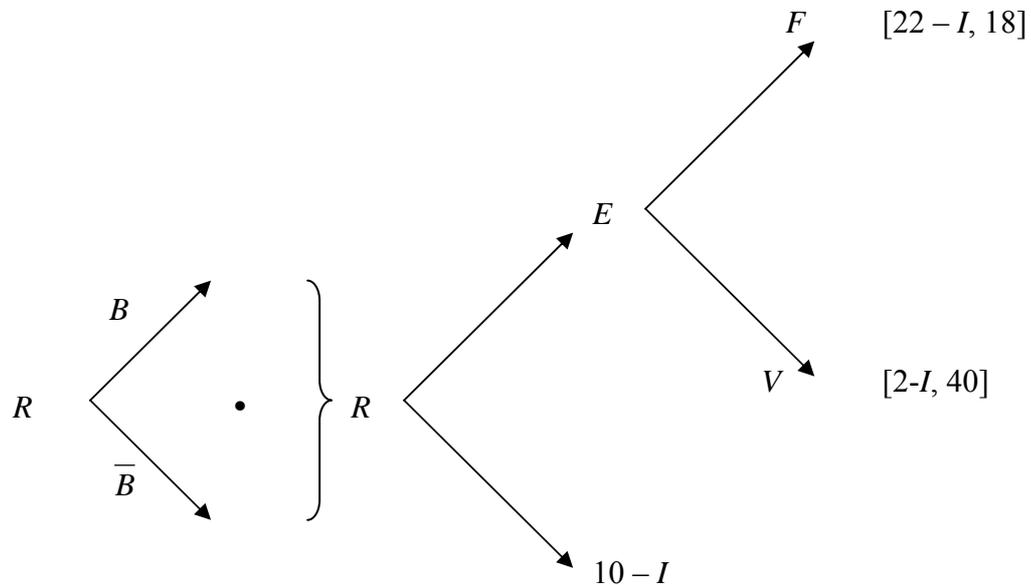
Trust is an *ad personam* belief or a belief we have about certain *types* of persons – e.g. devout people, the Ghurka, older women, one’s team mates. And “in virtually all [basic trust] games which occur naturally, the truster sees or otherwise observes the trustee before deciding. She therefore can, and should, use these observations as evidence for the trustee’s having, or lacking, trustworthy-making qualities” (Bacharach and Gambetta 2001: 148). In real life we know something about our partners, and often rationally *want* to know something about them. Thus, while still allowing R to make their choice blind if they so wish, we will give some information on E to R in order to approximate this reality.

We thus propose to add the following additional elements to the experiment protocol.

1. Before making his choice, R is offered the choice to buy a package of information concerning salient attributes of E (such as age, political preferences, membership of associations, ethnic origins) at a cost of I. We tell subjects in advance what the information is about, and administer three *information treatments*: ‘low’ and ‘high’ cost information, as well as no opportunity to buy information.

2. Whether or not R buys the information, R then decides whether or not he wishes to make the loan of  $M^*$

To focus on the information treatments, we restrict  $k=2$ . The new payoff structure is illustrated in the following chart:



for  $I = 0, 1, 2$

In words, if R does not make the loan but chooses to pay for information, his payoff is his endowment *less*  $I$ ; if R makes the loan, then his payoff is reduced by  $I$ ;

The different treatments in this extended experiment are: (a) no information available; (b) no charge for information; (c) ‘low’ information cost; and (d) ‘high’ information cost. The ‘free of charge’ information treatment (b) permits the unbiased estimation of the effects of information on trust. Otherwise, the estimated effects may be subject to a selection bias, for those who choose to buy may differ in critical ways from those who choose not to buy information.

*Key Features and open problems of Extended Design*

Before we say what information we plan to use, we should say that in general we assume that if R chooses not to buy information, his ‘blind’ trust or distrust is related to some prior expectation he has on the likelihood of trustworthiness of the population of Es to pay loans back. By contrast, if R chooses to buy the particular information that we offer, this means that

- his priors are not clear enough to make him either trust or distrust;
- that he believes the information we offer, about which he will know, may help him to make a more considered decision;
- he is prepared to sacrifice some of his winnings before deciding.

Questions following the implementation of the experiment will be addressed to those who do not buy information to ascertain whether they do not buy because they have their priors clear and set in stone, or because they simply think that the information that we offer is useless. We would also ask those who purchased information, which pieces they found most useful, and all will be asked what additional information they would have liked to have.

We expect that whether or not R buys will also depend on the cost of the information we set in the different treatments. We also expect that the set of Rs who do not buy information – those who are ‘blind-trusters or distrusters – will be different types from all other Rs who do purchase information. They will nurture a more extreme propensity with two horns: they will be inclined to trust and to distrust without worrying too much. Their dispositions will be more firmly etched and unconditional than among those who buy information. We expect the analysis of the attributes of

these types to reveal how they differ on a number of individual attributes. Finally, it seems likely that among those who do not buy information we will find more blind-*distrusters* than blind trusters. We hypothesise that a distrustful attitude will be more likely to be coupled with unconditional priors than a trustful attitude.

For the choice of *how many items* of information to use we have the constraint that we cannot introduce too many treatments, and we will probably choose a maximum of *three* binary variables.

For the choice of *which items* of information to use we have two options – one is to use attributes about which we suspect or have some evidence that people treat as trust- or distrust-warranting properties, whether rightly or wrongly. Co-ethnicity is one of them, gender is another. And at this stage we also want to keep other features in play, such as political preferences and religious beliefs. We can use survey evidence for this in which people are asked whether they trust certain categories more than others.

The other option is to use categories that we know affect trustworthiness regardless of whether people know that they do. We have little information on the latter – apart from age being positively correlated, which is found in many studies, and which we will definitely use. Gender is more uncertain, even though many people expect women to be more trustworthy, in this kind of game at least.

So it may be best to use the first option and find out ex-post whether they are correct predictors of people trustworthiness or they are prejudices, and which of them R value. To acquire this knowledge we will administer an ex-post questionnaire.

One problem with giving information about E is that the revealed attributes may induce other responses not related to an assessment of E's trustworthiness, but which increase or decrease the willingness to cooperate with that type of person. For instance, R with left leaning preferences may dislike, regardless of trust, dealing with E with opposite political views; or R can feel inclined to pass on their money to an E who is an old lady because he thinks she is more likely to be needy. The problem of generosity as a motive that we are trying to remove may thus come back through the back door. Again, we would ask ex-post in the questionnaire, which of the categories made you more inclined to give/not to give, and which made you more confident that E will pay you back.

### *Implementation*

We aim to collect a larger sample, of about 1000 subjects, for this second experiment, the FBTG with information. Each R is randomly allocated to an 'attribute group' of E's, defined by the set of attributes of E (e.g. age and sex) that will be 'for sale'. R chooses whether to buy information or not at cost I, and then chooses whether to make the loan or keep the endowment net of information costs. *Within an attribute group*, E is randomly matched with an R *who has made a loan* allocated to that group.

### **Conclusion**

One disadvantage of our design is that we will not be able to compare our results with the standard non-binary TGE, which are the most common ones. Furthermore, we may not be able to generalise to trust situations other than a financial situation like the one framed, but that is in the nature of the problem of trust decisions and not circumvented by the standard TGE.

We believe, furthermore, that the advantages discussed earlier far outweigh the disadvantages. An additional analytical advantage is that our binary trust game makes it possible to consider more attributes because we do not have to cover a range of values of amounts sent, as would be the case with the conventional TGE.

The experiment combined with the representative sample, and the latter's considerable size, should allow us to obtain good and representative measures, and to analyse the correlations between these and the individual variables collected by the survey questionnaire for these subjects. We anticipate that this analysis, of which there is virtually no previous example, especially as far as trustworthiness is concerned, will yield a wealth of results that will significantly enhance our understanding of the determinants of trust and trustworthiness.

Finally, let us recapitulate our hypotheses. In general, we expect deviation from self-interested rationality: i.e. some Rs will trust and some Es will fulfil trust. Attributes X...Z of R will affect their propensity to trust. In particular, we expect Rs whose responses to the risk questions identify them as being more risk-averse will be less likely to trust. Attributes X...Z of E will affect their propensity to fulfil trust. There are additional hypotheses related to particular treatments.

Special effects:

R will be less likely to trust in the 'all or nothing treatment'. E will be more likely to fulfil trust in the 'all or nothing treatment'.

Information purchase:

The information bought by R will affect their trusting decisions relative to otherwise identical R who were not given the opportunity to buy it. The information potentially knowable by R will affect E's probability of fulfilling trust. R will respond to the information cost when deciding whether to buy information. Attribute X...Z of R will affect their propensity to buy information.

## Bibliography

- Bacharach and D. Gambetta. 2001. Trust in signs. In *Trust in Society*, ed. K. Cook 148-184. New York: Russell Sage Foundation.
- Bacharach et al. 2001. Is trust self-fulfilling? An experimental study. Department of Economics Discussion Paper No. 76, University of Oxford.
- Barr, A. 2003. Trust and expected trustworthiness: experimental evidence from Zimbabwean villages, *Economic Journal*, 614-630.
- Bellman, C. and S. Kroeger. 2003. On representative trust. CentER Discussion Paper No. 2003-47, Tilburg University.
- Berg, J. et al. 1995. Trust, reciprocity, and social history, *Games and Economic Behavior*, 10:122-142.
- Bohnet, I. and R. Zeckhauser. 2004. Trust, risk and betrayal, *Journal of Economic Behavior and Organization*, 55-467-484.
- Bohnet, I. and S. Huck. 2004. Repitition and reputation: implications for trust and trustworthiness when institutions change, *American Economic Review*, 94: 362-366.
- Bohnet, I. et al. 2005. Learning trust, *Journal of the European Economic Association*, 3:322-329.
- Camerer, C. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton: Princeton University Press.
- Camerer et. al. 2005. Neuroeconomics: How neuroscience can inform economics, *Journal of Economic Literature*, 43:9-64.
- Coleman, J. 1990. *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Dasgupta, P. 1988. Trust as a commodity. In *Trust: Making and Breaking Cooperative Relations*, ed. D. Gambetta, 49-72. Oxford: Basil Blackwell.
- Dohmen T. et. al. 2005. Individual risk attitudes: new evidence from a large, representative, experimentally-validated survey, Discussion Paper 511, DIW Berlin.
- Fehr, E. et. al. 2002. A Nation-Wide Laboratory, *Schmollers Jahrbuch*, 122:519-542.
- Fehr, E. and B. Rockenbach. 2003. Detrimental effects of sanctions on human altruism, *Nature*, 422: 13 March.
- Gambetta, D. and H. Hamill. 2005. *Streetwise. How Taxi Drivers Establish their Customers' Trustworthiness*. New York: Russell Sage Foundation.

- Glaeser E. et al. 2000. Measuring trust, *Quarterly Journal of Economics*, 115:811-846.
- Jäckle, Annette, Emanuela Sala, Stephen P. Jenkins and Peter Lynn. 2004. Validation of Survey Data on Income and Employment: the ISMIE Experience, Working Papers of the Institute for Social and Economic Research, paper 2004-14. Colchester: University of Essex.
- Snijders, C. and G. Keren. 1999. Determinants of trust. In *Games and Human Behavior. Essays in Honor of Amnon Rapoport*, eds. D.V.Budescu, I.Erev, and R.Swick. Mahwah, NJ: Lawrence Erlbaum Associates, 355-383.
- Zak, P.J. Forthcoming. The neuroeconomics of trust. In *Two minds. Intuition and Analysis in the History of Economic Thought*, ed. R. Frantz. London: Springer.

## **Appendix 1 – BHPS sample**

It is our intention to carry out the first stage of our experimental study on a sample of about 200 individuals drawn from the ‘low income’ sub-sample of the UK part of the European Community Household Panel Survey (ECHP). This sample was interviewed annually from 1994 to 2001 – on eight occasions in total. Since 1997 it was administered by the Institute for Social and Economic Research (ISER) and undertaken jointly with the British Household Panel Survey (BHPS) activities. The sub-sample was selected in 1997 from the ECHP Great Britain panel. Selection was based on characteristics associated with low income – direct income data was not available at that time. Households were eligible if all adult members had been interviewed in the previous wave, and one of the following applied: household reference person unemployed currently or in the last year; household reference person receiving lone parent benefit; rented housing; receipt of means-tested welfare benefits.

Funding for the ECHP expired in 2001, but respondents who participated in the survey in 2001 were interviewed once more in early 2003 for purely methodological purposes (see Jäckle et al 2004). Interviews were sought with all ECHP sub-sample members who responded at the 2001 BHPS, that is, with 1,163 individuals in 781 households. New entrants and those not interviewed at the previous wave were not interviewed; eligible movers were followed to their new address. The interviews were based on the 2002 BHPS questionnaires. Efforts were made to maintain the context of the interview as far as possible, although both the household and individual questionnaires were shortened by removing some sections or questions not needed for the project. In particular the sections on demographics, health and values and opinions

were reduced considerably. Other instruments used in the 2002 BHPS, such as the self-completion questionnaire, youth questionnaire, telephone and proxy interviews, were also omitted. Interviews were completed with 88.8% of the 1,163 eligible adults. Non-interviews were due to untraced moves (2.8%), refusals either by the individual (3.2%) or the entire household (3.5%), or other non-coded reasons (1.6%).

The 2003 sample is not representative of the GB resident population in 2003, because low-income households are over-represented. Additionally, the sample is affected by attrition between 1994 and 2003. But the sample still contains a wide range of population groups, and the existence of a number of earlier waves of data for this sample provides a very rich background of information on these individuals. As an indication of the over-representation of poorer people, only 27% of the 2001 ECHP low-income respondent sample had bought the house they lived in, compared to an estimated 68 % of the 2001 census population. Similarly, the majority (52%) did not have a car for their private use, compared to 27% of the population. Also, a lower proportion of the survey sample is in employment (49%) than in the census population (60%), with a larger percent being retired (20% cf. 14%). In addition, only 10% of the sample have a degree or equivalent qualification (compared to 20% of the population) and 43% have no qualifications such as A-levels, O-levels or CSEs (cf. 30% of the population).

To quote Jäckle et al (2004), their 2003 sample “over-represents the older population and under-represents younger age groups compared to population estimates. Employment rates and educational qualifications are lower in the survey sample, as are the proportions of owner-occupiers and consumption indicators such as the

number of cars in the household. As to the causes of these differences, the age distribution is clearly affected by attrition: younger sample members are less likely to complete the interview. The survey also seems disproportionately to lose men, individuals with intermediate qualifications (A or O-levels, CSE), self-employed or unemployed, and those receiving zero benefits. For other employment and benefit characteristics the differences between respondents and non-respondents are not significant. Differences compared to the population estimates might therefore be attributed solely to the initial sampling process. However, the small number of employees may be masking differences caused by attrition.” By the time we carry out our experiment, the bias toward older people will be worse.

## Appendix 2 – Interview script

### Contact

The first step involves selecting the Rs names from the sample mentioned in Appendix 1. At least two methods for selecting the sample are possible: random sampling for the 2003 sample, thereby carrying over the biases from the 2003 sample, or stratified sample (e.g. by age) that attempts to obtain a more representative experimental sample.

Rs are contacted following the same procedure used by BHPS to contact its subjects (e.g. with BHPS headed paper), so as to take advantage of the trust relation subjects have with BHPS.

In the letter they are asked:

*We are running a nation-wide social experiment on how people make simple financial decisions, and your name, has been randomly selected as a possible participant. The experiment, carried out by an interviewer, will take only 20 minutes. You will receive £10 [£12] to participate in the experiment. Depending on your decisions, you will have the opportunity increase the amount of £10 [£12] that we give you to £22 [£24]. Your decisions will not be known to the interviewer and your identity will as usual remain strictly confidential.*

*We will call you in the next few days to answer any question you may have about the experiment and to ask whether you accept.*

*If you accept, we will agree a convenient time when the interviewer can call by your house.*

If they refuse, new letters will be sent to replacement names.

## PART I

### Interview of Rs

#### Stage 1 – Pre-interview

If he/she agrees to participate, R receives amount of £10 [£12] as compensation for taking part in the interview *cum* experiment. The money, in the form of a cheque [or two cheques, for £10 and £2], is sent to them before the interviewer visits. R is requested not to cash the cheque[s] before the interviewer visit. A letter sent to the participants contains the following:

*You have kindly agreed to participate in the social experiment that we mentioned in our earlier letter. Please find enclosed a cheque for £10 [two cheques for £10 and £2] to compensate you for your time. The cheque(s) are post-dated; they can be cashed 4 weeks after the interviewer visits. You will have the opportunity to increase the amount of £10 [£12] that we give you to £22 [£24]. However it is possible that depending on what you and others in the experiment decide, you could end up with nothing [£2].*

#### Stage 2 – Decisions explained

At the visit, the interviewer says:

*You have received £10 [£12]. Now we ask you to make a choice. You can choose whether to keep it, or to make a loan of £10 to someone we have paired you with.*

*If you keep it then our experiment ends here. You can cash the cheque(s) we sent you in four weeks.*

*If you loan it, then we will multiply it 4 times so the other person will get £40.*

*He or she will then be asked to decide whether to return the loaned sum to you with interest, in which case you will receive £22 [plus the £2 cheque you have kept], or whether to keep all the money, in which case you receive nothing [other than the £2 cheque you have kept]. You will find about the other person's decision and receive payment, if any, in about four weeks.*

*In this experiment we let everyone freely decide what they want to do, so there is no punishment if he/she decides to keep the money and return nothing to you.*

The interviewer then hands over a summary card with the payoffs, as follows

If you keep the money, you gain  
your £10 [£12]

If you loan £10 you can  
Either receive £22 [£24]  
Or nothing [£2]

[NB We need two cards per interviewer one for each treatment, namely one for standard exposure and one for high exposure]

### **Stage 3 – Purchase of information [only relevant for the second experiment]**

Once reassured that R understands the game, the interviewer will so proceed:

*Before you decide whether to keep or loan your money, we want to offer you the opportunity to acquire some information about your potential partner. We can tell you their age category, their gender and their X, Y, Z. This will cost you I.*

*What would you like to do? Buy the information or decide without it?*

If you buy the information and then decide to keep the money you gain your £10-I.

If you loan you can  
 Either receive £22-I  
 Or £2-I

If the subject buys, the interviewer will show them the card of the E that has been matched with the subject, else the interviewer moves to stage 4

Your potential partner is

A man/woman  
 Above/below 40  
 X,Y,Z

#### Stage 4 – Decisions

The interviewer hands a card to the subject along with a piece of paper that looks just like the £10 cheque sent previously and envelope 1. The interviewer offers to leave the room while the respondent fills out the card and places either the cheque or the piece of paper (phoney cheque) in the envelope and seals it. On the card there is

[subject code number]

*Tick your decision*

$\Delta$  I keep all the money

*Insert card and the phoney cheque in the envelope and seal.*

$\Delta$  I make a loan of £10

*Insert card and the real cheque for £10 in the envelope and seal.*

**Stage 5 -- Questionnaire**

Once the subject has decided and inserted the card in the envelope and sealed it the interviewer says:

*Lastly we would like to ask you a few questions concerning your decision on this sheet, which you will then insert in this envelope.*

The interviewer hands envelope 2 and the following sheet

[subject code number]

Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on the scale, where 0 means “unwilling to take risks” and the value 10 means “fully prepared to take risks”.

Are you generally a person who is fully prepared to take risks in trusting strangers or do you try to avoid taking such risks? Please tick a box on the scale, where 0 means “unwilling to take risks in trusting strangers” and the value 10 means “fully prepared to take risks in trusting strangers”.

*[for second experiment only]*

If you have purchased information about your partner can you please say which effect each item had on your final decision of whether to loan the money or not?

None    Some    Crucial

Sex

Age

X

Y

Z

If you have not purchased information, why not?

*Insert sheet in envelope and seal*

**Stage 6 – Conclusion**

After collecting the sealed envelope the interviewer says:

*We will process your decision and, if you loaned £10 and your partner returns your loan, we will send you a cheque for £22 in due course. In any case, we will notify you about the outcome. This should take about four weeks.*

## **PART II**

### **Interview of Es**

Following the selection of Es names from our sample, Es are contacted following the same procedure used by BHPS to contact its subjects, so as to take advantage of the trust relation subjects have with BHPS.

In the letter they are asked:

*We are running a nation-wide social experiment on how people make some simple financial decisions, and your name, has been randomly selected from the BHPS sample as a possible participant. The experiment, carried out by an interviewer, will take only 20 minutes. Depending on your decisions, you will have the opportunity to earn £18 or £40. Your decisions will not be known to the interviewer and your identity will as usual remain strictly confidential.*

*We will call you in the next few days to answer any question you may have about the experiment and to ask whether you accept.*

*If you accept, we will agree with you a convenient time when the interviewer can call by your house.*

If they refuse, new letters will be sent to replacement names in the same attribute group. In this case we have to have reserve names within attribute groups.

### **Stage 1 – The decisions explained**

The interviewer explains:

*In this experiment we have matched you with another person. This person received £10 [£12] from us for participation in the experiment.*

*We then gave this person the choice to either keep all the money or loan £10 to you.*

*He or she decided to loan it to you knowing that in this case we would increase the sum to £40, which is the sum you now have at your disposal. This person does not know your identity. I myself have had no contact with this person, who was interviewed by a colleague.*

#### **FOR FBTG WITH INFORMATION EXPERIMENT ONLY**

*[We offered to tell this person that you are MALE/FEMALE etc.*

*EITHER he/she declined*

*OR accepted our offer and paid I to know that you are MALE/FEMALE etc.]*

*But I stress that he/she does not know who you are. This person does not know your*

*identity. I myself have had not contact with this person, who was interviewed by a colleague.]*

*Also you must make your decision in private.*

*We now ask you to decide whether you want to keep all the money or return £22 to the person who made you the loan and keep only the rest.*

*You must write your answer on this card and put it in the envelope*

*Let me stress that you are absolutely free to choose either option, without fear of any consequence*

[subject code number]

*Tick your decision*

$\Delta$  I keep the money

$\Delta$  I return £22 and keep the rest

*Insert card in envelope and seal*

#### **Stage 4 -- Questionnaire**

Once the subject has decided and inserted the card in the envelope the interviewer says:

*Lastly, we would like to ask you a few questions on this sheet, which you will then insert in the envelope.*

[subject code number]

The person who gave you the loan had some information about you. Which effect do you think each item had on his or her decision of whether to loan the money or not?

	None	Some	Crucial	Don't know
Your Sex				
Your Age				
X				
Y				
Z				

*Insert sheet in envelope and seal*

### **Stage 5 – Conclusion**

After collecting the sealed envelopes the interviewer says:

*We will process your decision, and will send you a cheque in a couple of weeks. Thank you for your collaboration.*