# The long shadow of income on trustworthiness

John Ermisch
Institute for Social and Economic Research
University of Essex

Diego Gambetta
Nuffield College
University of Oxford

INSTITUTE FOR SOCIAL
& ECONOMIC RESEARCH

**Non-technical summary**

We employ a behavioural measure of trustworthiness obtained from an experiment carried out with a sample of the general British population whose individuals were extensively interviewed on earlier occasions. These previous interviews allow us to have very good income measures, and in particular to construct a measure of relative income that uses past income as a reference point. Our basic finding is that given past income, higher current income *increases* trustworthiness and, given current income, higher past income *reduces* trustworthiness. Past income determines the level of financial aspirations and whether or not these are fulfilled by the level of current income affects trustworthiness. But past income has a disproportionately large effect on trustworthiness compared to that predicted by the relative income theory, and this leads us to suspect that past income may also capture heterogeneity in relevant subjects' dispositions, with more opportunistic subjects being less trustworthy *and* having higher average incomes. We suggest and estimate a two-tier model in which relative income has the same positive effect within each past income class, but people in higher past income classes have a lower fundamental levels of trustworthiness.

The analysis also has interesting dynamic implications: trustworthiness should increase (decline) during an economic boom (recession) because current income rises (falls) relative to past income; in other words, trustworthiness is pro-cyclical. In so far as trustworthiness positively affects economic growth, then periods of recession would be deeper and longer as a result of trustworthiness being pro-cyclical. It would further follow that redistributive policies, while politically and financially easier during periods of boom, would in fact be more effective during recessions.

In terms of secular income growth, as long as it is not associated with a change in the distribution of the underlying traits that affect both income and trustworthiness it should have no long-term effects on the latter. But if that distribution is not immutable but interacts with changes in cultural zeitgeist, it would follow that, for example, the spreading of a more mercantile culture could increase the frequency of opportunism. Their income would grow faster as a result, but their fundamental level of trustworthiness would decline—justifying the often heard common sense perception that as societies become richer they become meaner.

# The Long Shadow of Income on Trustworthiness

John Ermisch[a] and Diego Gambetta[b]

Abstract

We employ a behavioural measure of trustworthiness obtained from an experiment carried out with a sample of the general British population whose individuals were extensively interviewed on earlier occasions. Our basic finding is that given past income, higher current income *increases* trustworthiness and, given current income, higher past income *reduces* trustworthiness. Past income determines the level of financial aspirations and whether or not these are fulfilled by the level of current income affects trustworthiness. We also suspect that past income may also capture heterogeneity in relevant subjects' dispositions, with more opportunistic subjects being less trustworthy *and* having higher average incomes.

**Key words:** trustworthiness, relative income
**JEL classification:** C93, D10

[a] Institute for Economic and Social Research, University of Essex and IZA; corresponding author.
[b] Nuffield College, Oxford University

*Introduction*

The division of labour that characterises modern societies would not be viable or as efficient without a considerable level of trust in strangers. In the words of Paul Seabright (2003),

> A division of labor needs to be robust against opportunism—the behavior of those who seek to benefit from the efforts of others without contributing anything themselves. In other words, participants need to be able to trust each other—especially those they do not know.

Being able to "trust each other" can only be sustained if strangers are trustworthy sufficiently often.

Yet, the mechanisms that affect trustworthiness remain to a large extent unknown. The growth of the scholarly literature on trust in the last quarter century has yielded a corpus of research, which while now reaching gigantic proportions, is enfeebled by having one foot of clay. Of the two choice components of trust, deciding whether to trust and deciding whether to be trustworthy, it is the former that gets the lion's share of attention, whether by studies which rely on surveys or on experiments.

While it has been common to try to measure levels of trust through survey questions, attitudinal questions about a person's trustworthiness would not, for obvious reasons, engender credible answers. The ample trust literature that relies on survey data is thus inevitably silent on trustworthiness and, as a consequence, also on the extent to which people's trust in others is a response to actual trustworthiness.

An alternative approach is to measure trust and trustworthiness through experiments that use real monetary rewards (see Camerer 2003, p.83ff., for a review).

---

Typical 'trust questions' include the following (from the World Values Survey, General Social Survey (US) and British Household Panel Study): 'Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?' In a recent paper, Ermisch et al. (2009)

This method has the advantage of providing behavioural measures and of being potentially clearer about the type of situation in which trust is an issue, the stakes and the reference group that subjects have in mind when making their experimental decisions—usually anonymous subjects in the experiment. Laboratory experiments, however, typically lack a sufficient range and variation of information on individual attributes needed to investigate the individual-level correlates of trustworthiness. In addition, they are often administered to students, usually self-selected subjects, unrepresentative of the adult population and who possess important knowledge about the "stranger" with whom they are interacting—he/she is another student.

In this paper, we try to overcome these limitations by combining the experimental method *and* the survey method, and by using a sample of the general population whose individuals were extensively interviewed on earlier occasions and about whose attributes much is thus known. This approach makes it possible to obtain a *behavioural* measure of trustworthiness, and, crucially, a rich and large data set of individual attributes. We employ a new experimental design of the so-called "trust game" developed by Ermisch and Gambetta (2006) that differs in a number of ways from the trust game used in most trust experiments. We believe that the differences we introduce allow it to provide better measures of the concepts of trust and trustworthiness.

---

question the usefulness of this question for measuring trust by showing it is uncorrelated with behaviour in the 'trust game' described below in this paper.

We know of only two other instances, neither of which was in the UK, in which trust-game experiments have been carried out with a representative sample of the population and combined with survey questions – one was carried out in Germany (Fehr *et al.* 2002) and gave us the original inspiration, the other in the Netherlands (Bellemare and Kroeger 2007). Other trust studies have also used population samples (e.g. Barr 2003 and Schechter 2007), but ones that are less representative of the population than these two studies.

Our main aim is to understand how trustworthiness between strangers varies within the population. Systematic variations in trustworthiness related to individual's attributes may, in turn, yield insights into what might influence a person's decision to fulfil a stranger's trust in them. We will focus in particular on the effects of income on trustworthiness.

The paper proceeds as follows. First we clarify what we mean by 'trust' and 'trustworthiness' and what might influence the latter, focusing in particular on people's economic resources. We then explain the experiment that we carry out and the key features of our design. Next we discuss our sampling frame and survey methods. The seventh section presents the main outcomes of the experiment, and the following three sections examine the main questions of the paper: how do economic resources affect trustworthiness in trust situations involving strangers. The final section presents our conclusions.

*Trust*

We work with a notion of trust that relates trust to specific acts and makes it easy to capture it empirically (Bacharach and Gambetta 2001). We trust when *we trust that someone will do X* – e.g. repay a loan, pay the fare, perform a given task as expected. The trust that we have in someone doing X does not necessarily extend to trust in that same person doing Y. We say that a person 'trusts someone to do X' if she acts on the *expectation* that he will do X when both truster and trustee know that two conditions obtain:

(i) if he fails to do X she would have done better to act otherwise–'if I knew he was a cheat I would not have lent him the money'. If trust is fulfilled the truster is better off than had she not trusted, but if trust is not fulfilled she is worse off;

(ii) her acting in the way she does gives him the opportunity to pursue a selfish reason not to do X – 'if I hadn't lent him the money he could not have cheated me'.

A *trustworthy* trustee is simply one who does X when those two conditions obtain.

In this regard , unlike the version we use (see below), the standard version of the trust game experiments (TGE) is laundered down to such a generic characterisation as to expunge the core expectation necessary to define a situation as one in which trust is involved.

*Trustworthiness*

In all TGE, including our version, trust and trustworthiness are not addressed in general – an impossible task – but refer to a financial 'game' in which players can earn real money: 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 there is overwhelming evidence, in real life as well as in experiments, that a considerable number of players trust and fulfil trust 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 framework we use to understand trustworthiness derives from Bacharach and Gambetta, 2001).

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 whenever it pays off; 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 forces, these trust-warranting properties (TWP) that can tame self-interest? There is a plethora of candidates. There are 'global' TWP which sustain trustworthiness in a wide range of encounters: these include fairness, honesty, reciprocity and trust-responsiveness, the sense of obligation to fulfil the trust that is put in one. And even people who are neither fair minded nor reciprocal can still have reasons to be trustworthy and stick to their promises. They may fear post-mortem divine retribution, or be socialised to respect contracts and act in a manner to promote good will in others. Furthermore, trustworthiness can be an eminently 'local' response that depends on what one is being trusted about. 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 TWP which sustains trustworthiness only in relationships which involve children's welfare.

Three conclusions can be drawn from the above. First, a trust experiment should try to elicit as wide a range of TWP as that which can make a trustee trustworthy in that situation in real life. In this regard the standard TGE has limitations. It is designed in such a way as to tap only into reciprocity or fairness as the properties that can lead the second

---

Triggering reciprocity and trust-responsiveness relies on the ability or motivation to read the intentions. Laboratory experiments by McCabe et al. (2003), using binary trust games similar to that used here, in which the truster's expectation is known to the trustee, suggest that the trustee's attribution of *intentions* to the truster is important in the trustee's decision to fulfil trust or not. The evidence in McCabe *et al.* (2003)

mover to return enough money to the first mover, which while important are not the only ones, not even perhaps the main ones.

Next, since in any one trust situation more than one property can be motivating the trustworthy action and different set of properties can be active in different situations, seeking an overarching theory of trustworthiness, even if only in financial transactions, is not a feasible or even meaningful aim.

Lastly, TWP share one feature: they are not directly observable. What we (and trusters) can observe are the individual attributes and behaviours that may activate or hinder trust-warranting properties. In this paper we focus on the TWP that could be associated to trustees' income, past and present, absolute and relative.

*Trustworthiness and income*

In trust game experiments, like our own, which measure trustworthiness as people's willingness to part with money, which they could keep with no detrimental consequences to themselves, it seems reasonable to think that their income, real or perceived, may come into their decision. Yet, little experimental research addresses this question; most

---

is not consistent with dominance of 'outcome-based preferences', such as preferences exhibiting 'inequity aversion' (Fehr and Schmidt 1999), which are assumed to capture ideas of 'fairness'.

In the TGE, if the first mover (truster) chooses to trust it is because he *trusts* the second mover (trustee) to make the transfer worth his while in the sense of expecting that trustees 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'.
More generally, in much of the experimental literature trustworthiness is conflated with some of the forces that may bring it about, but these forces are not coextensive with it. Reciprocity may often occur without involving trustworthiness: 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 think of his reciprocation as an act of trustworthiness. 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 or he simply likes to do what he *knows* is expected of him. The key to a trust situation minimally defined is that the trustee knows that the truster *expect*s him to do X, in the conditions we describe in the text.

research, while measuring trustworthiness, does not have data on income, and amongst that which does, almost none, somewhat oddly, focuses on this question. Moreover, the evidence is scarce, inconclusive and some of it based on designs unsuitable for this question. Some research finds no effect, other a positive effect and other still a negative one. Our data from, both from the experiment and from the linked panel survey, are to our knowledge the richest available to address this question.

One could think, first of all, that what matters is the *absolute* level of income. This it could do in two opposite directions. The simplest idea would be that people who have more resources feel more disposed to be trustworthy in a financial exchange—the same cost of being trustworthy weighs proportionally less on their finances than it does on those of poorer people. For instance, taxi drivers make such an assumption when they display a preference for picking up richer customers whom they deem unlikely to do a 'runner' on them to avoid paying the taxi fare (Gambetta and Hamill 2005). Alternatively, one could expect the opposite effect to dominate: richer people are richer because they pursue their self-interest more vigorously, with less intrinsic regard for others, while poorer people are poorer because they are more inclined to share and take a collective approach to welfare rather than an individualistic one.

---

Bellemare and Kröger (2007) and et al. (2002), who use a similar method to ours, in 'no variables excluded' regressions find *no effects* for current income whether on trust or trustworthiness. Barr and Serneels (2006), in a sample of 164 workers from a number of Ghanaian firms, find a *positive* correlation between second mover's earnings and odds of returning more than the amount sent by the first mover (both players are in the same firm and they know it). Kropp et al. (2009) find that "while the absolute level of family income had no significant effect on repayment behaviour, US borrowers that perceive themselves as having a family income that was relatively lower than other US households repaid at higher rate." (p.67). Bonein and Serra (2009: 784) in a sample of 148 undergraduates find that a measure of disposable income (which they do not describe) is *negatively* correlated with the odds of returning something in a continuous trust game. A negative correlation is also found by Becchetti and Conzo (2009, Table 7). They use the sum of each family member's disposable income as measure, and the data on trustworthiness from a trust game experiment with 152 microfinance borrowers in Argentina and a control group of non-borrowers of the same size. Although household income is only one of two determinants of trustees' decisions to be significant in the regression, they do not even discuss the finding.

Instead of absolute levels of income, however, a subtler expectation would be that it is the *relative* income that matters. It is not whether one is rich or poor, but whether one *feels* richer or poorer that makes the difference. In this paper we will investigate relative income with reference to the subject's past income, and not with reference to the income of others (as in Kropp et al. 2009), which our design does not allow us to measure.

According to the psychological theory of adaptation (or 'habituation'), we get used to a stimulus as it persists over time, but we are responsive to *changes* in circumstances, including financial ones.   The effect of changes in income could in turn be interpreted as having two effects on trustworthiness. One is related to the notion of reference points in 'prospect theory'; for instance, those who have recently experienced an improvement in their financial situation are above their reference point, and being trustworthy is perceived as cheaper to them.   Alternatively, they may feel richer as a result of the improvement and this may cause them to act more benevolently and generously towards others. In the former case the driving mechanism would be a cost calculation biased downwardly by the experience of a financial improvement, in the latter it would be a positive emotional state induced by the improvement.  But in practice it is difficult to distinguish between the two because perceived lower costs make you feel better off as well.

*The experiment*

The game we presented to subjects, whom we shall call the truster (R, a 'she' for convenience) and the trustee (E, a 'he'), is a basic 'one-shot' trust game in binary form.

---

We are grateful to Abigail Barr for drawing our attention to this possibility.
  This is related to the concept of 'income comparisons', in this a comparison to one's past self rather than to others.

R receives a £10 banknote at the onset of a professional interviewer's visit to her home; it is described as compensation for taking part in the interview *cum* experiment. R is told that she will have the opportunity to obtain £22 if she gives the money to another person (E) with whom she has been randomly matched and about whom she is told nothing. She is told that the experimenter will increase it by £30, and so E will receive £40, and that E will then be given the choice to either pay back £22 to R, or keep all £40. R is informed that if she decides to give £10, she will know the outcome in about four weeks, and if E decides to pay back the specified amount she will receive a cheque for £22. The procedures ensure that the interviewer does not observe or otherwise know her decision (for further details see Ermisch et a. 2009). If R keeps the £10, the game ends there.

If, however, R chooses to pass the £10 on, an E enters the picture, and he is the focus of the analysis in this paper. E is offered, by a different interviewer who visits him at his home, a binary choice of whether to pay back £22 and keep the difference or keep the whole £40 he received. E is told that R was informed that E would be making this choice before R decided to pass on £10. E is given two cheques made out in his name: one for £40 pounds and one for £18, either of which he can cash without delay. Again, the interviewer does not know his decision. Our procedures are thus, in a sense, 'double blind': the interviewer knows the subjects' name and address but not their decisions, while the researcher knows the subjects' decisions, but identifies them by their code numbers, not associating them to their name or address.

---

Of course £22 in four weeks is not necessarily worth the same to a person as £22 now. But for £10 now to be worth the present value of £22 in four weeks, R's *monthly* discount rate would need to be 120%, or an annual rate in excess of 1400%. In any case, this is not relevant for the focus of this paper on the trustworthiness decision.

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 him (e.g. see Bacharach and Gambetta 2001, Bohnet and Huck 2004 and Bohnet et al. 2005).

For one-half of the sample we used a second treatment, in which R receives £12 in cash (a £10 note and a £2 coin) as compensation for taking part in the interview *cum* experiment. R is offered the binary choice of either keeping the whole £12 or giving £10 to E and keeping £2. The remaining steps are the same as in the previous treatment. The outcome variable of interest is:the probability that E pays back the £22 ('fulfils trust').

*Key features of our design*

Our design differs from that of the standard trust-game experiment (TGE) (Berg et. al. 1995, Glaeser et al. 2000), also used by Fehr et al. (2002) and Bellemare and Kroeger (2007). Ermisch and Gambetta (2006) provide a detailed critique of the standard TGE, arguing that the standard TGE lacks the basic features of a trust situation even in a one-shot case, and, as a result, it blurs the link with clear notions of trust and trustworthiness and introduces confounding effects.

In particular, the binary TGE outlined above more precisely captures a clear notion of trustworthiness. This is because in the standard TGE, E can return any amount to R. In our design, R knows in advance how much she can expect back from E if E fulfils trust, and E knows that this is an expectation of R when she makes her decision. 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 (e.g. making and repaying a loan). We

believe that our design strongly encourages E to put himself in the situation of a person who has benefited from R's expectation that he will return £22. The study by Karlan (2005), of poor Peruvian women, participating in a micro-credit program, indicated that Es in the standard TGE who returned a larger proportion are less likely to default on their loan from the program, suggesting that the trust game behaviour indeed reflects individual-specific trustworthiness.

In so far as beliefs are based on experience, past experience is important for people's decisions in experiments. It is precisely participants drawing on their social context that should give our experiment external validity. We believe that our design makes it highly likely that trustees' decisions reflect their habitual practices in these situations—their sense of reciprocity, fairness or sensitivity to obligations—and that the decisions of trusters reflect their expectations of such behaviour.

As noted above, we introduce two payoff structures. We expect that the treatment in which R has the draconian option of parting with all her money or keeping it all, will *increase* the chances that E will pay his due back if R passes the money on. 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: E will feel that R does not suffer a total loss of face and that he is not leaving her with a worrisome and salient *nothing*.

*The Sample and Survey Methods*

The sample frame was households who were formerly members of the British Household Panel Survey (BHPS); they were an over-sample of poorer people recruited in 1997 and dropped from the panel for technical and funding reasons in 2001 and re-interviewed in

2003 for a special study (Jäckle et al 2004). We randomly selected one person from each household. An advance letter asking these persons whether they would agree to participate explained that we would be 'running an experiment on how people make simple financial decisions' as well as asking questions from a short questionnaire similar to earlier ones. 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. Also, these subjects are used to receiving compensation (a voucher) for their participation. The experiment and subsequent interviews were carried out face-to-face by professional interviewers at the subject's home. Interviewers were instructed to read only from the experimental script and not to elaborate further. We randomly matched E-players to the Rs who passed on the money.

The experiment was carried out first. The two treatments were applied to one-half of the sampled Rs each. After completing the experiment, the respondents filled out a short confidential self-completion questionnaire on their own that followed up their

---

As a consequence of using the same survey organisation that carries out the BHPS, many of the respondents knew the interviewer from previous contact through the BHPS. For example, 63% had the same interviewer as they had in 2003, the last year they were interviewed. On the one hand, this is an advantage because it reinforces respondents' trust that the experiment will be carried out as described e.g. Eckel and Wilson report that a large proportion of subjects participating in a one-shot trust game did not believe that they were matched with a real person (2004; p.458, fn.9)). On the other hand, subjects may be more predisposed to trust and be trustworthy because they know the interviewer, despite the fact that we reiterated a number of times in the interview script that the interviewer would not know the subject's decision. We carried out some tests (Ermisch et al. 2009), and even though we cannot rule out entirely the possibility that familiarity with the interviewer may encourage people to be more trusting and trustworthy, none of the correlations was statistically significant.

decision. They then completed a short questionnaire with the interviewer (CAPI) and finally another confidential self-completion questionnaire. On average, the experiment took 10 minutes and the remainder of the survey about 15 minutes.

In the first stage (i.e. Rs), 173 of 245 eligible cases were interviewed, a response rate of 71%. In the second stage (Es), 85 of 127 eligible cases were interviewed: a response rate of 67%. Almost all of the non-response arose because of inability to locate the selected individual who had moved house. Since we know that the probability of moving in a particular year is inversely related to length of current residence (e.g. Morrison (1971) for an early study and Belot and Ermisch (2009) for evidence from the BHPS), we checked whether non-response may have affected our inferences by relating the odds of giving the £10 (among Rs) and returning the £22 (among Es) to length of residence. We found no significant effects, even after controlling for other variables that may affect trust or trustworthiness. Thus, it appears that those who move more often do not behave differently in terms of trust and trustworthiness.

Because of its origin, our sample differs from the British population in a number of potentially important dimensions: it over-represents women, people who are retired, older, divorced or separated, and those with lower incomes. At a minimum, our sample can be taken as representative of households with low to moderate income. The focus on lower income people may have its serendipitous advantages because the sums of money involved in the experiment, particularly the £40 that Es receive, may be more important to them.

*Basic Results*

Overall, we find that about two-fifths of people are willing to trust strangers by giving them £10 in the expectation of receiving £22 if the stranger is trustworthy. One-half of strangers fulfil their trust, resisting the temptation of keeping £40 rather than only £18.

**Table 1: All or Nothing Treatment Effect**

| Behaviour | All or Nothing Treatment | £12 Treatment | *p-value* *N* |
|---|---|---|---|
| % of Rs giving £10 | 39.5% | 45.0% | 0.48 |
| N | 81 | 80 | |
| % of Es returning £22 | 59.5% | 44.7% | 0.19 |
| N | 42 | 38 | |

A somewhat larger percentage of Es return £22 in the all or nothing treatment as we expected (Table 1), although the difference is not statistically significant at conventional levels. Still, we cannot rule out that the thought of R being left with £2 inclines E to be untrustworthy. This could explain the puzzling finding of many trust experiments whereby in many cases the second mover sends only a little money back, but not enough to cover the sum sent over by the first mover – arguably a guilt allaying tip rather than a clear trustworthy act.

*Income, current and past*

What does motivate one-half of these subjects to resist the temptation to pocket £40 and return £22 to the trusters? The first of two hypotheses we have is that the current absolute level of resources affects the trustworthiness (TW) decision either positively or negatively. To test for this we use a banded *household* income question that was asked in

the survey following our experiment.    We have ten bands, the lowest is less than £150 a

week and the highest is more than £1150 (2007). In all the intermediate bands we use the

mid-point value. For the two extreme bands we took from the 2007 BHPS the average

value of income for people in those bands (£90 and £1615 respectively).  Quartiles of the

distribution of the current household income from the experimental sample are shown by

the solid bars in Figure 1, while the horizontally striped bars show the corresponding

quartiles for a representative sample from the 2007 BHPS.  The much lower level of

income in the experimental sample reflects its origins as an over-sample of poorer people.

**Figure 1: Quartiles of the Distributions of Current and Average Past Monthly
Household Income from Experimental Sample (n=75) Compared with Distribution
of Household Income from 2007 BHPS (n=7598, weighted), 2007 prices**



---
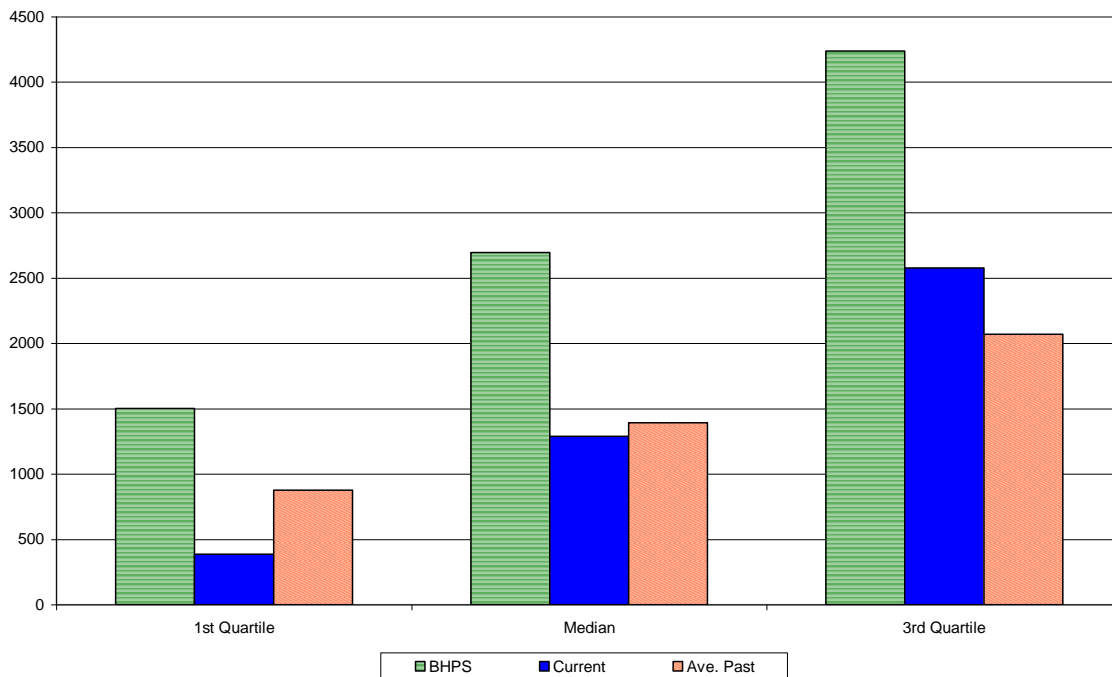
We also examined the impact of net wealth in 2001, subjective description of current (2007) financial
well-being and home ownership on return behaviour in the experiment, but found virtually no effect of
these.

As we can see from column 1 in Table 2, current household income is estimated to have virtually no effect on TW (because one-half of our sample is trustworthy, the logit coefficients can be transformed to effects on the probability of TW around the average by multiplying by 0.25). This and the other results described are robust to the introduction of control variables – sex, age, highest qualification (3 categories), ease of understanding the experiment (interviewer's evaluation), good health and marital status. But the current income measure could be considered to be a noisy measure of income. The respondent is asked to provide an estimate of his entire household's income and place it in one of ten bands, which we then translate to midpoints of the income band. If the measurement errors are of the classic type, then our estimate of the income effect is biased toward zero.

**Table 2: Impacts of current and past income on Trustworthiness[a]**

| Return £22 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Current household income (log) | 0.12 (0.24) | -- | | | |
| Average log household income 1997-2001 | | -1.08* (0.46) | -1.36* (0.49) | -1.30* (0.53) | -1.31* (0.49) |
| Father's Occupational Prestige (Hope-Goldthorpe score)** | | | 0.044* (0.013) | 0.045* (0.014) | -- |
| Years of residence in current house | | | | 0.008 (0.023) | -- |
| Difference between own and father's HG score** | | | | -- | -0.007 (0.011) |
| Log-likelihood | -52.559 | -53.649 | -50.877 | -50.807 | -53.011 |
| Number of parameters | 2 | 2 | 4 | 5 | 4 |
| Number of observations | 76 | 82 | 82 | 82 | 82 |

[a] Standard errors in parentheses.
*Significant at 0.05 level or less

As alternative to the banded income question, we use an individual's subjective assessment of their financial situation arranged along 4 categories. While the results show that the more comfortable the financial situation the higher TW, but taken as a group the categories are not significant. This is also the case when we include the measures of past income and father's occupation discussed below.

**Also includes a variable indicating missing information for the father, which is never statistically significant.

We have however an arguably better measure of income which we can use for a further test of possible effects of absolute income on TW. It is based on *past income* in the years 1997-2001 (our experiment was carried out in 2007, and we lack income information for the interim years); we obtain this measure by exploiting the sampling frame described earlier, which provides 5 years of panel information for the individuals in our sample. Each individual in the household was asked to report income in the past month from a large number of sources (e.g. earnings, benefits, etc.). Not only are these reports elicited by detailed questions, each individual's own income is aggregated into household income in a consistent manner. We take the average of log household income during this five-year period, which should remove much of the transitory fluctuations in past income and reduce measurement error, particularly desirable given our small sample. The diagonally striped bars in Figure 1 show the distribution of this measure of income transformed to be in terms of pounds sterling at 2007 prices (Appendix Table 1 shows descriptive statistics for all income measures in logs) . Compared to current income, the distribution of average past income is more compressed, as we would expect; its standard deviation is 962 compared to 1559 for current income. The correlation coefficient between current and past average income is 0.46 (0.56 in terms of their logs). Column 2

---

Some caution should be exercised in comparing the level of average past income and current income in the experimental sample. The former has been adjusted to 2007 prices by using the retail price index, but it may not be a good reflection of the prices of the basket of goods consumed by our low income sample. As we work in logs of the income variables, this does not matter for the subsequent analysis.

of Table 2 shows that the past average measure of income is strongly, significantly and negatively related to TW in the experiment.

How can we interpret this effect, especially vis a vis the lack of an estimated effect of our measure of current income? Because of its more detailed measurement and averaging over five years, past average income could be a better measure of the fundamental level of income than the measure of current income we have in the sense that it has a higher signal-to-noise ratio.

While this may explain the lack of an effect for current income it does not solve the other puzzle: why would income levels in the somewhat distant past cast such a remarkably strong and *negative* shadow on TW? Let us explore two conjectures. In both cases, we hypothesise that being or feeling more dependent on others' help to achieve desired outcomes increases TW. But in one case this 'dependency' reflects differences in the class experiences of subjects and in the other 'dependency' reflects heterogeneity in subjects' traits, which in turn affect their class and income.

*Class?*

Our first conjecture is that the past income measure could capture the effects of the material conditions that affect people's lives in ways that shape their pro-social behaviours like trustworthiness. Four studies reported in Piff et al (2010) test the hypothesis that 'lower class' people are more concerned with the needs of others than 'upper class' people, and so will act in a more pro-social manner to improve others' welfare. The reasoning is that lower class (poorer) people experience less personal

---

We obtain a similar estimate when we use the log of mean past income rather than the mean of the log of past income.

control and are (or feel) more dependent on others' help and good will to achieve desired outcomes compared with upper class (richer) people, who are characterised by economic independence, more personal control and more freedom of personal choice. As a consequence, lower class people are motivated to behave in ways that promote social engagement with others, making them more accustomed to behave pro-socially in social interactions like the one in our experiment. Thus, this could explain why persons with higher past average income are *less* likely to be trustworthy in our experiment.

To test for this interpretation we can consider the impact of 'class' on TW directly. One measure of class, not influenced by 'type', is the subject's father's 'social class or occupation.' A continuous measure related to this, available in our data, is the father's occupational prestige score based on the calibration suggested by Hope and Goldthorpe (1974). This score is also highly correlated with earnings. Column 3 of Table 2 shows that people whose fathers were in higher status occupations are *more* likely to be TW, in contrast to the argument of Piff et al., and the negative impact of past income increases. The person's own Hope-Goldthorpe score also increases TW; its impact is statistically insignificant, but the impact of the father's score remains about the same (results not shown).

On the whole these results do not support the 'class' interpretation of the negative effect of past income on TW, operating through dependency on others.

*Heterogeneity?*

Our second conjecture is that the past income measure captures differences in the disposition towards trustworthiness, in that fundamentally richer people would be 'more opportunistic', and thus capable of ending up with higher incomes, but also, for the same

reason, less intrinsically trustworthy. That is, they may well be trustworthy in an environment in which trustworthiness is enforced, but also capable of taking advantage of situations in which the pursuit of their raw self-interest remains anonymous and has no negative consequences, as in the experiment.

If these differences are driving the negative relationship between past income and TW, then we should find a corresponding relationship between TW and other indicators of 'opportunism'. One such indicator could be the frequency of residence changes, which would be higher the more intense is the search for better opportunities and the weaker are the social ties with others. Belot and Ermisch (2009), for instance, show that people who have a lower propensity to make close friends move house more often. We do not have a measure of residence changes but can use a proxy: we know that the probability of moving in a particular year is inversely related to length of current residence (e.g. Morrison (1971) and Belot and Ermisch (2009)). We would thus expect that TW should increase with length of residence. Column 3 of Table 2 shows almost no evidence of such an association, and the impacts of past income and father's occupation change little.

Another indicator of opportunism could be the degree of social mobility: people who are upwardly mobile are more opportunistic and vice versa, and so we might expect that people whose own Hope-Goldthorpe score is higher than their father's are less TW. In column 5 of Table 2, we include the difference between own and father's HG score as a regressor rather than father's HG score. Consistent with our reasoning, this variable is negatively associated with TW, but the association is small and it is mainly driven by the positive impact of father's occupational status demonstrated in the other analyses.

These results are inconclusive with respect to this conjecture, and our puzzle of why past income shows such a strong negative effect on TW stands intact. Before we come back to it, however, what about our other hypothesis, that income matters for trustworthiness in relative terms?

*Relative income*

We estimate the impacts of two notions of 'relative income' on trustworthiness, which differ in their reference points. For one of them the reference point is income in the very recent past. To investigate this we use the answers to a question we have in the questionnaire which subjects completed after the experiment, which records whether they thought, that in terms of their financial situation, they were better off, worse off or about the same as they were the previous year. When included in the logit model, these categories do not discriminate TW in a statistically significant way (p-value=0.47, results not shown), either when included on their own or in the other models discussed below.

In contrast to the recent past as reference point, we now take *past average income* (in the years 1997-2001) as reference. This measure is closely akin to the notion of relative income in the theory of Richard Easterlin (1973, 1974). He postulated that a person's expected standard or living, or 'material aspirations' for short, is based on his past experience. A person compares his current income to his aspirations, and when the former is higher relative to the latter he feels better off, and, in his main applications, is more disposed to marry and have children. Past average income may thus reflect

---

The relative income concept is also used to explain other behaviours, such as labour supply, savings and reported levels of happiness; in particular, the notion is invoked to account for the lack of an upward trend in people's reports of 'happiness' in surveys, despite considerable growth in incomes—the so called Easterlin paradox (see Clark et al (2008) for a review).

'material aspirations' *à la* Easterlin and, as such, it may be one component of relative income.

**Table 3: Change in Household Income between 1997-2001 and 2007 and Trustworthiness**

|  | Change in Household Income Relative to Average | | |  |
|---|---|---|---|---|
|  | Decline>30% | −30%≤Change≤30% | Increase>30% |  |
| Percent | 34.7 | 37.3 | 28.0 | 100 |
| *N* | 26 | 28 | 21 | 75 |
| Percent returning £22 | 38.5 | 53.6 | 66.7 | 52.0 |

Table 3 shows the distribution of household income changes between 1997-2001 and 2007 relative to the average change in real household income. It indicates that those who had the largest decline in real household income exhibit the lowest chances of returning the £22 in the experiment. A simple comparison between those who experienced a change in income greater than the average (56 percent) relative to the rest shows a significantly larger percentage returning £22: 64 percent cf. 36 percent (p-value=0.016). Thus, we have some evidence that relative income directly affects trustworthiness.

Given the small size of our sample, it is however more efficient to use the continuous measures of current and past average income rather than these discrete measures. We define the log of 'relative income' as the difference between log current and past average log household income. From column 1 of Table 4 it is apparent that relative income has a strong positive effect on TW. We tested whether increases and declines in current income relative to past income had the same effects and found virtually no difference.

## Table 4: Impacts of current and past income on Trustworthiness[a]

| Return £ 22 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Relative household income (log) | 1.09* (0.29) | | | | | 1.25* (0.30) | |
| Current household income (log) | | 1.25* (0.33) | 1.09* (0.34) | 1.32* (0.41) | 1.30* (0.32) | | 1.25* (0.32) |
| Average log household income 1997-2001 | | -2.74* (0.75) | -2.78* (0.77) | -2.83* (0.82) | -2.79* (0.83) | | -3.36* (1.33) |
| Father's Occupational Prestige (Hope-Goldthorpe score)** | | | 0.034* (0.016) | 0.036* (0.017) | -- | | |
| Years of residence in current house | | | | 0.035 (0.023) | -- | | |
| Difference between own and father's HG score** | | | | -- | -0.021 (0.015) | | |
| Middle Tercile, Average log household income 1997-2001 | | | | | | -1.28* (0.64) | -0.10 (0.59) |
| Top Tercile, Average log household income 1997-2001 | | | | | | -1.62* (0.68) | -0.81 (1.14) |
| Log-likelihood | -47.441 | -42.953 | -41.523 | -40.436 | -42.278 | -43.756 | -42.420 |
| Number of parameters | 2 | 3 | 5 | 6 | 5 | 4 | 5 |
| Number of observations | 75 | 75 | 75 | 75 | 75 | 75 | 75 |

[a]Standard errors in parentheses
*Significant at 0.05 level or less
**Also includes a variable indicating missing information for the father, which is never statistically significant.

Provided that the reference point (aspirations) is proportional to past income, the relative income hypothesis has a rather sharp prediction: equal and opposite effects of current and past log income. Of course, if the reference point is a more general function

---

If we use the log of average past household income rather than the average of the logs in our measure of relative income the coefficient (std. error) is 1.09 (0.29), i.e. identical to two decimal places.

More specifically, our econometric specification is the following: $TW_t = \alpha + \beta \ln(y_t) + \gamma \sum_{t-k} \ln(y_{t-k}) + e$, where $TW_t$ is a latent variable for trustworthiness in the experiment period $t$, $y_t$ is household income in year $t$, $e$ has a logistic distribution, the sum is over $k$ from 5 to 10 and $w_{t-k}$ are weights that sum to unity. We assume $w_{t-k} = 1/5$ for $k=5,..,10$. Our relative income hypothesis is that $\beta = -\gamma$. In terms of Easterlin's ideas, each past year's contribution to 'aspirations' is proportional to income in that year and the constant of proportionality is absorbed in the constant term, $\alpha$. Alternatively, we could assume that 'aspirations' are proportional to average past income ($\sum y_{t-k}/5$), and again the proportionality factor is absorbed in the constant term. The estimates of $\beta$ and $\gamma$ are very similar in this alternative specification because the correlation coefficient between average log past income and log average past income is 0.94.

of past income, then the sharp prediction disappears.  But then any set of coefficients on current and past income that has the former's positive and the latter's negative is consistent with the relative income hypothesis.  The results are not consistent with the sharp prediction.  In column 2 of Table 4, we include current and past log incomes individually in the logit model for TW, and find that current income does show a considerable effect which it did not have when introduced on its own; but past income still has a much stronger proportionate effect on TW.   The difference between the coefficients of current and past log income in column 1 of Table 4 is -1.49 (std. error=0.59), which is statistically significant ($p$=0.012).

One reason that the size of the effects of current and past log income differ may be because, as explained earlier, current income is a noisier measure.  If this were to explain fully the difference in coefficients and there were no noise in the average past income measure, the noise to signal ratio for log current income would need to be about 0.55 (i.e. the ratio of the measurement error variance to the true variance).  This seems implausibly large for the measurement issue alone to explain that difference.

If we choose to ignore the measurement issue, a more substantive reason could be that past income may be picking up other influences on TW, due to heterogeneity.  It is worth noting that, in contrast to its role in the relative income hypothesis, past long term income, under this hypothesis, would only be a 'marker' of heterogeneity with no causal relationship to TW; that is, opportunistic types both have higher incomes *and* are less TW.

The impacts of current and past income are robust to changes in specification like those in Table 2.  Columns 3, 4 and 5 of Table 4 add father's occupation, residential

tenure and the relative occupation levels of respondent and father, respectively. The coefficients of tenure and relative occupation are in the same direction as in Table 2 but are larger in size, although they are not statistically significant at conventional levels. Despite the fact that these variables may be related to 'opportunism', the introduction of them barely affects the coefficient of past income, leaving its disproportionately large effect unexplained.

*Discussion*

We propose a way of combining the 'opportunism' and relative income hypotheses in our empirical analysis: we take a person's broad position in the past average income distribution as the indicator of unobserved behavioural traits related to TW and average past income. Within each income class there could be a relative income effect.

Column 6 of Table 4 is an empirical representation of this conjecture. We divide long term income into terciles, and we see that in each income class (third), relative income has essentially the same effect—a test of the constraint that the effect of relative is the same in each class yields a chi-square statistic (2df) of only 0.06 (p-value=0.97). At the same time, the probability of TW falls as you go up the income distribution: going from the lowest to highest third, the probabilities of TW are 0.74, 0.44 and 0.36, respectively (evaluated at average relative income for the entire sample). So while being

---

Another possibility is that those who had higher incomes in the past than currently may be disproportionately those who retired during the intervening period and the movement into retirement may affect TW. When we include an indicator for retiring between 2001 and 2007 in the model of column 2 (8% of the sample), the coefficients of current and past log income are virtually identical, and the impact of retirement (std. error) is 0.72 (0.41).

The number of income classes that we can use is constrained by our small sample size. Statistical tests indicate that father' occupation, residential tenure and difference between father's and own occupation can be excluded from the model.

richer over the longer period goes hand in hand with being (a lot) *less* trustworthy, because it reflects being a more opportunistic type, higher relative income *within each class* of long term income still makes one *more* trustworthy. This is illustrated in Figure 2: the boxes represent a person's long-term income class and the effect of differences in relative income within a class is represented by the upward slopping lines within each box.

## Figure 2: Two Tier Model

**Trustworthiness**

**Relative income**

**Long-term income**
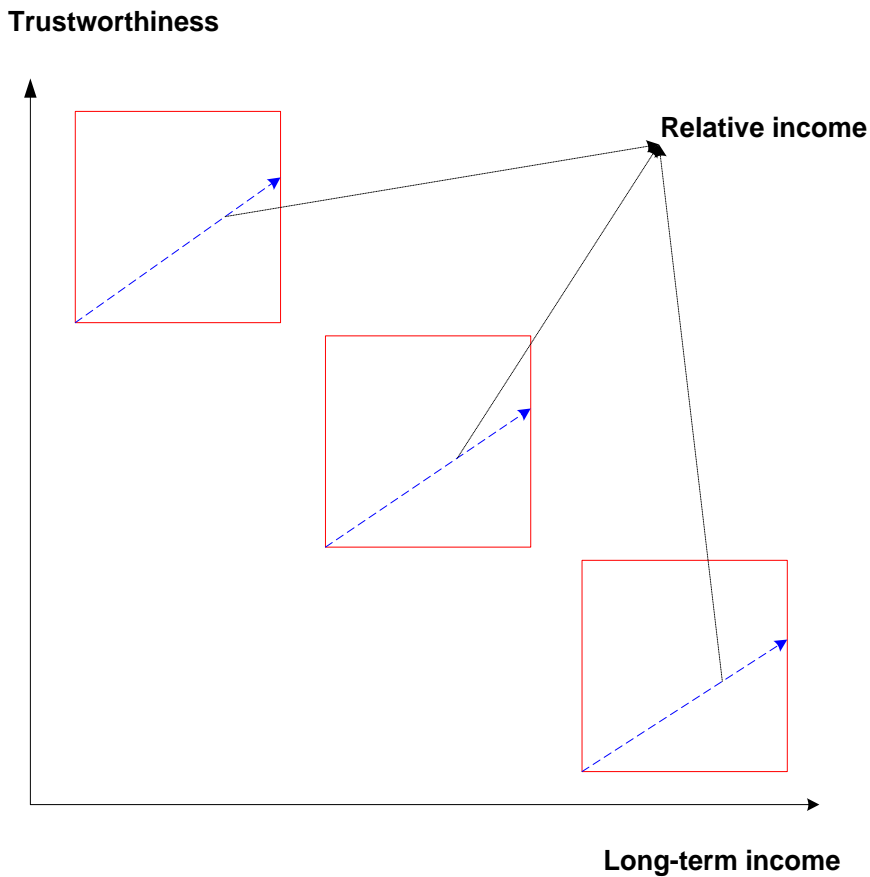
Does the model in column 6 of Table 4 do justice to the data? In column 7 we present estimates of a model which relaxes the relative income constraint on the impacts of current and past log income imposed in column 6. The hypothesis that the impacts have equal and opposite signs is rejected (chi-square (1df) =8.96; p-value=0.003). Indeed the tercile indicators are jointly insignificant, with their exclusion taking us back to model 2 in Table 4. While model 2 accounts for the data in a statistically more parsimonious way, it is however harder to interpret in substantive terms because past income may partly have a causal impact as a component of relative income, but also serve as a mere indicator for opportunism with no causal implication.

*Conclusions*

Our basic robust finding is that given past income, higher current income *increases* TW and, given current income, higher past income *reduces* TW. This suggests that the more distant past is relevant as a reference point for relative income and that higher relative income increases trustworthiness. Here the driving mechanism of TW would be whether aspirations set by the level of past income, are or are not satisfied by current income. But past income also has a disproportionately large effect on trustworthiness compared to that predicted by the relative income theory, and this leads us to suspect that it may also capture an heterogeneously distributed trait that affects both income and TW, 'opportunism', with more opportunistic types being less trustworthy and having higher average incomes. We suggest a two-tier model in which relative income has the same positive effect within each average income class, but people in higher average income classes have a lower fundamental level of trustworthiness. While the empirical

representation of this model is somewhat ad hoc—assuming three income types when the underlying variable may be continuous—and does not account for the data quite as well as a more parsimonious model, it has theoretical appeal.

It also has interesting dynamic implications: TW should increase (decline) during an economic boom (recession) because current income rises (falls) relative to past income; in other words, trustworthiness is pro-cyclical. In so far as trustworthiness positively affects economic growth, then periods of recession would be deeper and longer as a result of trustworthiness being pro-cyclical. It would further follow that redistributive policies, while politically and financially easier during periods of boom, would in fact be more effective during recessions.

In terms of secular income growth, as long as it is not associated with a change in the distribution of the underlying traits that affect both income and TW it should have no long-term effects on TW. But if that distribution is not immutable but interacts with changes in cultural zeitgeist, it would follow that, for example, the spreading of a more mercantile culture could increase the frequency of opportunism. Their income would grow faster as a result, but their fundamental level of trustworthiness would decline—justifying the often heard common sense perception that as societies become richer they become meaner.

---

We are grateful to Pedro dal Bò for pointing this out to us.

**Appendix Table 1: Log Income Measures (n=75)**

| Log        Income Measure | Mean | Std. Deviation | 1st Quartile | Median | 3rd Quartile |
|---|---|---|---|---|---|
| **2007 Current** | 7.08 | 0.86 | 5.96 | 7.16 | 7.85 |
| **Average log 1997-2001** | 7.21 | 0.53 | 6.78 | 7.24 | 7.64 |
| **log average 1997-2001** | 7.29 | 0.58 | 6.86 | 6.33 | 7.69 |

# References

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

Barr, A. 2003. Trust and expected trustworthiness: experimental evidence from Zimbabwean villages. *Economic Journal* 113, 614-630.

Barr, A. and Serneels, P. 2006. Reciprocity in the workplace. *Experimental Economics*, 12, 99-112.

Becchetti, L. and Conzo, P. 2009. Creditworthiness as a signal of trustworthiness. AICCON Working Papers #67-2009, www.aiccon.it/file/convdoc/wp_aiccon_ creditworthiness.pdf

Bellemare, C. and Kroeger, S. 2007. On representative social capital. European *Economic Review* 51, 183-202.

Belot, M. and Ermisch, J.F. 2009. Friendship ties and geographical mobility: evidence from the Great Britain. *Journal of the Royal Statistical Society*, 172:437-444.

Berg, J., Dickhaut, J. and McCabe, K. 1995. Trust, reciprocity, and social history. *Games and Economic Behavior* 10, 122-142.

Bohnet, I. and Huck, S. 2004. Repetition and reputation: implications for trust and trustworthiness when institutions change. *American Economic Review* 94, 362-366.

Bohnet, I., Harmgart, H., Huck, S. and Tyran, J.-R. 2005. Learning trust. *Journal of the European Economic Association* 3, 322-329.

Bonein, A. and Serra, D. 2009. Gender pairing bias in trustworthiness. *The Journal of Socio-Economics* 38 (2009) 779–789

Camerer, C. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction.* Princeton: Princeton University Press.

Clark, Andrew, Frijters, Paul and Shields, Michael A. 2008. Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature* 46:95-144.

Easterlin, Richard A. 1973. "Relative Economic Status and the American Fertility Swing." In *Family Economic Behavior,* ed. E. B. Sheldon. Philadelphia: Lippincott.

Easterlin, Richard A. 1974. Does economic growth improve the human lot? Some empirical evidence. In *Nations and Households in Economic Growth. Essays in Honor of Moses Abramovitz*, ed. R. David and M. Reder. New York: Academic Press.

Eckel, C. C. & R.K. Wilson, Rick K., 2004. "Is trust a risky decision?" Journal of Economic Behavior & Organization, Elsevier, vol. 55(4), pages 447-465

Ermisch, J.F. and Gambetta, D. 2006. People's Trust: the design of a survey-based experiment. Working Paper of Institute for Social and Economic Research, paper 2006-34 (PDF). Colchester: University of Essex.

Ermisch, John, Gambetta, Diego, Laurie, Heather, Siedler, Thomas and Uhrig, S.C. Noah. 2009. Measuring people's trust. *Journal of the Royal Statistical Society* Series A. 172:749-769.

Fehr, E. & Klaus M. Schmidt, 1999. A Theory Of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics*, MIT Press, vol. 114(3), pages 817-868.

Fehr, E., Fischerbach, U., von Rosenbladt, B., Schupp, J. and Wagner, G.G. 2002. A Nation-Wide Laboratory. Schmollers Jahrbuch 122, 519-542.

Gambetta D. and Hamill, H. 2005. *Streetwise. How taxi drivers establish their customers' trustworthiness*. New York: Russell Sage

Glaeser E.L., Laibson, D.I., Scheinkman, J.A. and Soutter, C.L. 2000. Measuring trust. *Quarterly Journal of Economics* 115, 811-846.

Goldthorpe, J.H. and Hope, K. 1974. *The Social Grading of Occupations: A New Approach and Scale,* Oxford.

Jäckle, A., Sala, E., Jenkins, S.P. and Lynn, P. 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.

Karlan, D. S. 2005. Using Experimental Economics to Measure Social Capital and Predict Financial Decisions. *American Economic Review* 95, 1688–1699

Kropp, J. D. , Turvey, C.G. , Just, D.R. ,Kong, R. and Guo, P. 2009. "Are the poor really more trustworthy? A micro-lending experiment", *Agricultural Finance Review*, Vol. 69 Iss: 1, pp.67 – 87.

McCabe. K.A, Rigdon, M.L. and Smith, V.L. 2003. Positive reciprocity and intentions in trust games. *Journal of Economic Behavior and Organization* 52, 267-275.

Morrison, P.A. 1971. Chronic movers and the future redistribution of population: a longitudinal analysis. *Demography* 8, 171-184.

Piff, Paul K., Kraus, Michael W.,  Cheng, Bonnie Hayden and Keltner, Dacher. 2010. Having less, giving more: the influence of social class on prosocial behavior, *Journal of Personality and Social Psychology*.  Advance online publication. Doi: 10.1037/a0020092.

Seabright, Paul. 2003. *The Company of Strangers*.  Princeton: Princeton University Press.

Schechter, L. 2007. Traditional trust measurement and the risk confound: an experiment in rural Paraguay. *Journal of Economic Behavior and Organization* 62, 272-292.