School of Economics Working Paper
2015-01

University of East Anglia

SCHOOL OF ECONOMICS

Honesty and beliefs about honesty in 15 countries

by David Hugh-Jones*

*School of Economics, University of East Anglia

Abstract

The honesty of resident nationals of 15 countries was measured in two experiments: reporting a coin flip with a reward for "heads", and an online quiz with the possibility of cheating. There are large differences in honesty across countries. Average honesty correlates with per capita GDP: this relationship is driven mostly by GDP differences arising before 1950, rather than by GDP growth since 1950, suggesting that the growth-honesty relationship was more important in earlier periods than today. The experiment also elicited participants' beliefs about honesty in different countries. Beliefs were not correlated with reality. Instead they appear to be driven by cognitive biases, including self-projection.

JEL classification codes

Keywords

School of Economics
University of East Anglia
Norwich Research Park
Norwich NR4 7TJ
United Kingdom
www.uea.ac.uk/economics

Honesty and beliefs about honesty in 15 countries

David Hugh-Jones*

September 25, 2015

Abstract

The honesty of resident nationals of 15 countries was measured in two experiments: reporting a coin flip with a reward for "heads", and an online quiz with the possibility of cheating. There are large differences in honesty across countries. Average honesty correlates with per capita GDP: this relationship is driven mostly by GDP differences arising before 1950, rather than by GDP growth since 1950, suggesting that the growth-honesty relationship was more important in earlier periods than today. The experiment also elicited participants' beliefs about honesty in different countries. Beliefs were not correlated with reality. Instead they appear to be driven by cognitive biases, including self-projection.

The 19th-century Chinese scholar Feng Guifen attributed British economic success to four causes: utilizing manpower, superior agriculture, control over rulers, and "the necessary accord of word with deed" (quoted in Spence 2001). These causes map nicely to standard answers to the biggest question in economics[1]: why are some countries richer than others? Feng's first two causes relate to the factors of land and labour, and the third to political economy explanations of economic growth (Acemoglu and Robinson, 2013). The fourth and final cause points to a cultural variable: the propensity to tell the truth and keep one's promises, that is, honesty.

Honesty might be good for growth because it encourages economic interactions beyond a narrow circle. For example, consider the Prisoner's Dilemma. Selfish rational players will defect. Altruists may prefer to cooperate. But even otherwise selfish players will cooperate if they make a mutual promise to do so, and prefer not to break their promises. When people are honest enough, bare promises act like contracts, allowing gains from cooperation in situations where formal contracts would be hard to write or enforce. If honesty varies across cultures, as Feng proposes, then this mechanism may explain why levels of wealth vary too.

On this account, honesty should be particularly important in environments with weak institutions (so that contracts are hard to enforce) and without technology for monitoring (so that contracts are hard to monitor).

*Department of Economics, University of East Anglia. D.Hugh-Jones@uea.ac.uk

[1]This widely-used phrase appears to come from McCloskey, McCloskey and Jr (1990).

That description fits most pre-industrial societies. Modern societies, on the other hand, have relatively effective institutions, plus technological advances that make monitoring easier. Cowen and Tabarrok (2015) list a number of recent technologies that improve monitoring and reduce moral hazard, including online reviews, credit checks, remote attendance monitoring for workers and escrow systems. Earlier advances include double entry accounting, the telegraph and telephone, and on the institutional side speedy access to justice and impartial bureaucracies. In these respects, even today's developing economies are far ahead of where they were in the past. So, we might expect honesty to be most important for economic development early in history.

This paper reports experimental research on cross-national differences in honesty. People from 15 countries took part in an online survey containing two incentivized experiments measuring honest behaviour. I use both the well-known coin flip experiment, where subjects report the result of a coin flip and are offered money for reporting "heads", and a new experimental paradigm: an online quiz in which subjects were able to cheat and this could be detected. Using two different experiments is important because subjects' behaviour may vary systematically across contexts. For example, in Japan, so for Japanese subjects the coin flip paradigm may prime specific associations which lead to a higher or lower level of dishonesty. Having two measures helps ensure that cross-country differences are not just different reactions to a particular experimental paradigm. I also use an online sample, rather than drawing from university subject pools. This sample is more diverse and closer to the population profile on some demographic variables, though still falling short of true statistical representativeness.

With data from the experiments, I first test Feng's basic intuition that levels of honesty differ between societies. Then, taking these country average levels of honesty, I relate them to economic development, and examine how this relationship changes in different historical periods. Lastly, I test whether honesty is related to a cultural variable that is often supposed to affect economic outcomes: religion.

Participants in the experiment were asked to predict the average honesty of other participants from different countries. The ability to realize gains from trade and cooperation may depend not just on partners' honesty, but also on their mutual trust. People hold beliefs about the honesty of both their fellow citizens, and those in other countries. Under rational expectations, these beliefs would be correct on average. However, a large literature in psychology examines the formation of *stereotypical* beliefs about groups (McGarty, Yzerbyt and Spears, 2002). Stereotypes may, but need not, be accurate (Jussim et al., 2009). The accuracy of different groups' beliefs about each other can affect how they interact. For instance, the German public's willingness to support debt bailouts may be affected by stereotypes about Southern Europeans (Soll, 2015). So it is important to understand how these beliefs are formed. The data support neither rational expectations, nor the alternative hypothesis that

people's derive their beliefs about others' honesty from their experience in their own country. Instead, these beliefs are driven by biases, including self-projection and, surprisingly, pessimism about the honesty of people in one's own country relative to those in other countries.

A growing literature in economics uses incentivized experiments to study honesty. Initial experiments examined whether people's unwillingness to lie was related to costs (Gneezy, 2005; Charness and Dufwenberg, 2006; Vanberg, 2008). More recent work has revealed variety in the motivations of liars: people may lie to benefit others (Erat and Gneezy, 2012) or in such a way as to make their lie less obvious (Fischbacher and Föllmi-Heusi, 2013). There is also some evidence for heterogeneity: women are less willing to tell a lie (Dreber and Johannesson, 2008). Abeler, Becker and Falk (2014) run the coinflip paradigm within a telephone survey of a representative sample of Germans, and find that complete honesty cannot be ruled out.

Within a broader literature comparing social preferences across cultures (e.g. Croson and Buchan 1999; Henrich 2004; Henrich et al. 2006; Herrmann, Thöni and Gächter 2008), some existing papers have compared honesty across countries. Holm and Kawagoe (2010) run a sender-receiver game in Sweden and Japan, finding no differences in average levels of truth-telling. Pascual-Ezama et al. (2015) report an experiment among university students in 16 countries. Subjects were asked to flip a coin and report the result; reporting heads was rewarded with a chocolate. Cross-country differences in honesty are small and insignificant. Mann et al. (2015) run a "dice rolling" game in five countries, with students and the public. In each of 20 trials, subjects privately chose to bet on the top or bottom face of a die, and were then shown the pips on these faces; if their chosen face scored less than the other, they could lie by reporting the other face. Payments were 10 US cents per pip. They find small, marginally significant differences in honesty. Both these studies offered relatively low incentives (a chocolate, and 20 cents average per lie, respectively). Dieckmann et al. (2015) run a "coin flip" game in five European countries and find significant differences in honesty between countries. They also examine subjects' incentivized assessments of different nations' honesty, with results similar to those reported here.

This paper also speaks to the broader literature on cross-country differences in trust and social capital. Most cross-country comparative research on trust uses survey data (Porta et al., 1997; Knack and Keefer, 1997). A potentially important proximate cause of trust is *trustworthiness*, including honesty. Survey questions on trustworthiness face the problem that respondents may not always answer truthfully;[2] behavioural data can therefore play an important role in measurement. Several papers have used the trust game (Berg, Dickhaut and McCabe, 1995) to measure trust and trustworthiness across cultures (Yamagishi, 1988; Buchan, Croson and Dawes, 2002; Holm and Danielson, 2005; Bornhorst et al., 2010). The paradigms used here have the advantage

[2]Hugh-Jones (2015) examines questionnaire measures of honesty in the context of this paper's experiments.

over the trust game of requiring no interactions between participants, which makes them simpler to administer to a sample across many countries. Indeed, no previous single experiment has run the trust game on more than a few countries.[3] A subtler difference is in what the experiments measure. Responder behaviour in the trust game may be affected by several factors, including altruism towards one's partner, reciprocity and a perceived social norm. In the experiments below, respondents face a relatively simpler choice between the honesty norm and material self-interest.

This paper's contribution, then, is to estimate the honesty of a demographically diverse sample in a large set of countries, using two different incentivized experiments; to relate the country-level results to economic development at different periods of history, and cultural background; and to examine how beliefs about honesty are formed.

Design

The experiment was implemented in Qualtrics, an online survey platform. Participants were recruited from 15 countries, in two waves. The first wave consisted of 804 respondents from Brazil, China, Greece, Japan, Russia, Switzerland, Turkey, and the United States. The second wave added 735 respondents from Argentina, Denmark, Great Britain, India, Portugal, South Africa, South Korea, and Turkey.[4] The countries were chosen to provide a mix of regions, levels of development, and levels of social trust. Initial analysis of the first wave of countries showed large cross-country differences. The second wave was then added so as to learn more about the correlates of honesty at country level. For the sake of simple exposition, this paper analyses data from both waves together, reporting statistics on the first wave only in a few places.

Participants were members of managed online panels. These are typically used by firms for market research. Members are recruited from across the web, and sign up to receive regular invitations to surveys and questionnaires. Recruitment materials usually emphasize both material benefits of taking part ("share your opinion to win gifts, cash and test products") and non-material benefits ("have your say"). Sample frames for individual surveys are then selected and a randomized sample of candidates is invited to participate by email. Sometimes these panel surveys are supplemented by ad-hoc recruitment directly from the web. To sample a particular demographic profile, quota sampling is used: individuals are invited until the required number of people from a given category (e.g. a particular gender and age group) has completed the survey, after which further individuals from that category are turned away. Thus, these panels are not probability samples of the country populations:

[3] Johnson and Mislin (2011) report a meta-analysis of different trust game experiments across 35 countries.

[4] Abbreviations: Argentina AR, Brazil BR, China CN, Denmark DK, Great Britain GB, Greece GR, India IN, Japan JP, Portugal PT, Russia RU, South Africa ZA, South Korea KR, Switzerland CH, Turkey TR, United States US.

quota sampling is used to provide balance across gender and age, but respondents are likely to have been unrepresentative in other dimensions. For example, all Swiss respondents answered in French, although German was available as a questionnaire language. This method falls short of the ideal of representative sampling, but it probably compares well with the method, widely used in cross-cultural experiments, of recruiting participants from laboratory subject pools in different countries. Quota sampling guarantees a reasonable spread of age groups, and other demographics are likely to be more similar to the country population than student samples. Another difference with a standard laboratory experiment is that no experimenter is present. This reduces control: for example, subjects may have answered in the presence of others. Conversely, it may also reduce experimenter demand and social desirability effects.

The experiment contained two different measures of honesty. In the COIN FLIP, respondents were asked to get a coin ready and had to confirm they had done this. On the next screen, they were asked to flip the coin and report the result. They were also informed that they would receive a money incentive (either $3 or $5) if they reported "heads". Thus, respondents who flipped "tails" faced a choice between telling the truth and receiving the money. This paradigm allow honesty to be estimated at aggregate level, from the difference between the proportion reporting heads in any group, and the 50% proportion expected if all report truthfully. It has been used in several previous experiments (e.g. Bucciol and Piovesan 2011; Fischbacher and Föllmi-Heusi 2013; Abeler, Becker and Falk 2014; Pascual-Ezama et al. 2015). In the QUIZ, subjects were given a test on the topic of music, consisting of six open-ended questions. They were asked not to look up the answers on the internet, and they had to tick a box confirming they had answered on their own before moving on. Since the survey was web-based, and respondents were not being monitored, they could always in fact look up the answers online. The questions were:

1. Who wrote the composition "Für Elise"?
2. What is Lady Gaga's real first name?
3. Name the drummer of the rock group Nirvana.
4. In what year was Claude Debussy born?
5. How many valves are there on a standard modern trumpet?
6. Name the town and state of the US where Michael Jackson was born.

The topic of music was chosen so as to minimize the test's cultural bias: Lady Gaga and Michael Jackson have global name recognition, and classical music is also known in non-Western societies. Questions 2, 4 and 6

were designed to be very difficult for almost anyone to answer, but very easy to look up online: they can be answered, in any of the quiz countries, by typing "Lady Gaga", "Debussy" or "Michael Jackson" into Google.[5] Participants were offered a money incentive ($3 or $5) if they answered all the questions correctly. Thus, they faced a temptation: they could cheat by looking the answers up, and lie by ticking the checkbox to affirm they had not done so.

The order of the coin flip and quiz experiments was randomized. After completing both experiments, subjects answered a 15 question integrity questionnaire taken from Whiteley (2012). Next, subjects were told that some people from one specific country out of the countries in their wave had also taken the coin flip question. They were asked to guess what percentage of people from that country answered "heads". A response of 50% would indicate a belief in the complete honesty of survey respondents from the named country; 100% would indicate a belief in their complete dishonesty. Half of the respondents were offered a $2 incentive if their answer was within 10 percentage points of the true figure for that country (treatment INCENTIVE). The experiment finished with a brief questionnaire on demographics. Wave 2 included additional demographic questions, and also added questions on self-reported unethical behaviour.

In total, the survey had four measures of honesty: two experiments, the integrity questionnaire, and the wave 2 questions on behaviour. Hugh-Jones (2015) reports the correlations between these measures in detail. In summary, behaviour in the quiz and coin flip experiments was significantly correlated across subjects, and also correlated with self-reports of past unethical actions, but answers to the integrity questionnaire were only very weakly correlated with the experiments. Neither kind of questionnaire data is used in this paper.

To avoid participants being influenced by the origin of the researchers (i.e., people may be less averse to lying to someone from a very different culture), there was no branding on the questionnaire. However, participants were told that data from their answers would be stored at the University of Essex, so they could have drawn inferences from this.

The sample within each nation was composed of equal numbers of men and women, and equal numbers of each of five age groups (18-24, 25-34, 35-44, 45-54, 55+). Payments were made to respondents by Qualtrics in addition to its standard payment to respondents and by the same mechanism: rewards are translated into points which can be spent on various items including vouchers, gift cards, tickets for cash lotteries *et cetera*. This payment mechanism, as well as purchasing power differences, makes it hard to equalize the value of incentives across countries. However, it is not clear whether and how lying behaviour responds to incentive size: this topic has been much debated in the literature. To check whether incentive sizes would affect my results, I randomized

[5] Google is not widely used in China, but other search engines give equally easy access to these answers.

the size of payments offered. Half the subjects were assigned to the HIGH payment treatment. They were offered $5 for reporting heads in the coin flip and $5 for getting all quiz questions right. The remaining subjects were assigned to the LOW payment treatment, where both payments were $3. This difference between payments is of the same order as the differences in a dollar's purchasing power between countries in the sample.[6]; if the HIGH/LOW treatments did not affect behaviour, this should lessen concerns that intercountry differences are driven by differences in purchasing power. The HIGH/LOW and INCENTIVE treatments, and the order of the two honesty experiments, were balanced within nations.

Results

Descriptive statistics are shown in the appendix. In this analysis, only country residents who stated that they were of the country's nationality are included (e.g. residents of Turkey who gave their nationality as Turkish).[7] Thus, numbers of respondents vary between countries.

Before looking at per-country results, I briefly report the association between the two experimental measures of honesty: the respondent's report on the coin flip, and his or her score out of 3 on the "hard" quiz questions 2, 4 and 6 – henceforth simply "quiz score". Median quiz scores were 2 for those reporting heads on the coin flip, and 0 for those reporting tails (Mann-Whitney p <0.001). Put another way, proportions reporting heads were 62% for those scoring 0 on the quiz, 65% for those scoring 1, 68% for those scoring 2, and 77% for those scoring 3. If the proportion reporting heads in a group is h, the group's estimated level of dishonest reporting is $2h-1$, so dishonesty about doubled from the lowest to the highest quiz score (24% vs. 54%).[8]

Result 1 *There was significant evidence for dishonesty in all countries.*[9]

The first graph in Figure 1 shows the proportion reporting heads by country. The unshaded part of the bars and the right hand scale show the country's estimated level of false reporting. In every country except Great Britain and South Africa, we can reject the null hypothesis that the proportion reporting heads is 50% at the 0.05 significance level (binomial test).

The second graph shows mean quiz scores by country The association between quiz scores and coin flip reports makes it plausible that both measures are capturing individuals' levels of dishonesty. Nevertheless, it

[6] A Big Mac cost between $1.83 (India) and $6.82 (Switzerland) among sample countries in July 2015 (The Economist, 2014).

[7] There are two exceptions: residents of the US and South Africa who gave their nationality as their ethnicity, e.g. "black", "white", "Hispanic", "Caucasian" etc; and UK residents who gave it as "English", "Scottish", or "Welsh".

[8] For further details on the link between the measures see Hugh-Jones (2015).

[9] All analyses use unweighted samples. Weighting age/gender groups, to reflect the size of these groups within each country's population, gives essentially unchanged results.

could be that subjects were answering these questions correctly without cheating. To check this, the quiz was administered to 144 students in a laboratory setting where cheating was difficult, with a £3 incentive for getting all six questions right.[10] Of the hard questions, while 25 out of 144 students got question 2 right, only two students got question 4 right, and only the same two students got question 6 right. Thus, a score above one on hard questions was extremely hard to achieve without cheating.[11] In each country, more than 30% of subjects scored above one; even if 10% of the population could achieve this without cheating, a rate much greater than the laboratory sample, the hypothesis of no cheating would be rejected in every country at $p < 0.001$.

[10] Participants in the laboratory were from 37 countries. The most common were Romania and the United Kingdom.

[11] Per-country mean quiz scores on the "hard" questions correlated at above 0.95 with three alternative measures: mean quiz scores on all questions, proportions getting more than one hard question right, and scores on questions 4 and 6 only.

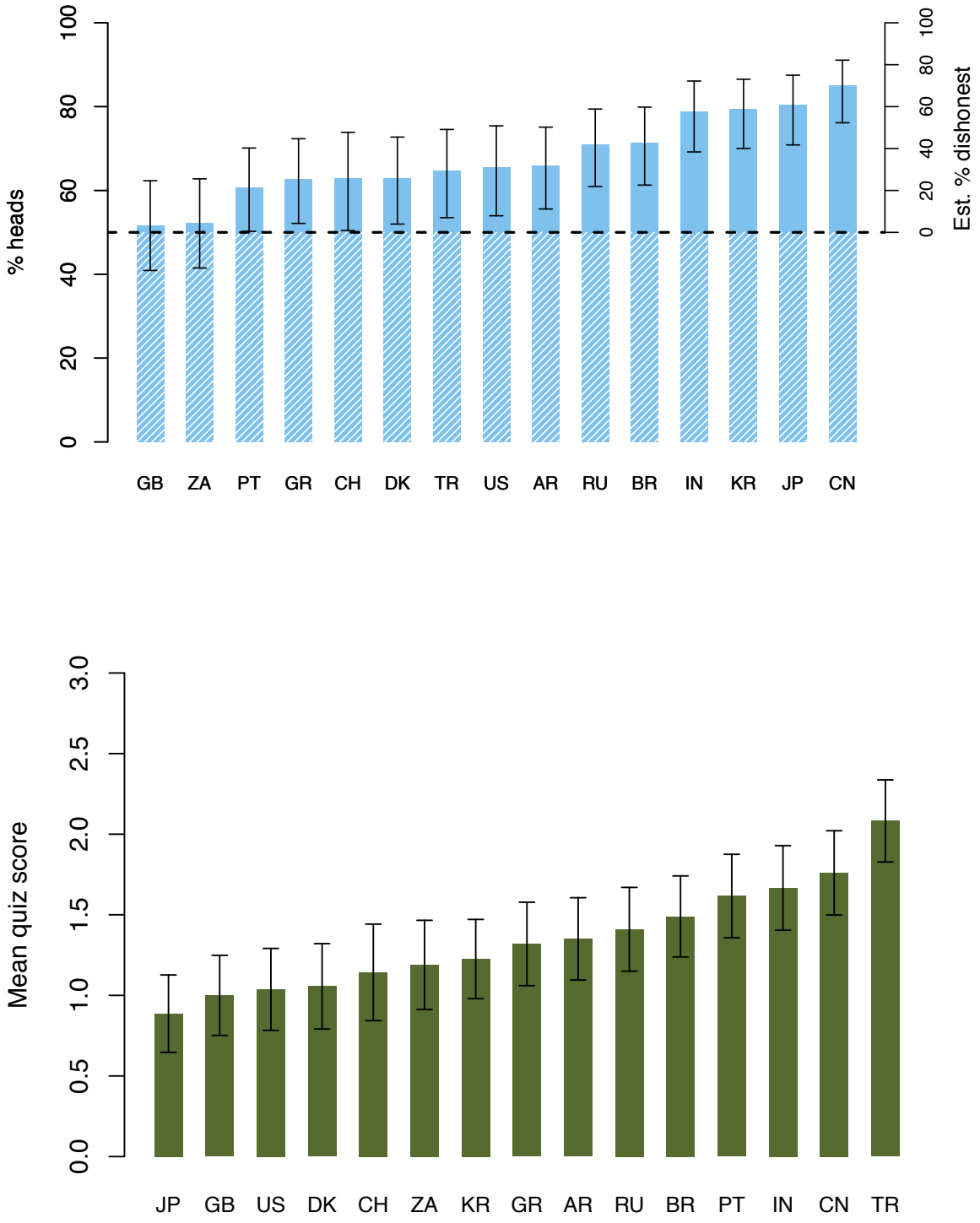

Figure 1: Experimental measures of honesty per country. Bars show 95% confidence intervals.

Result 2 *Levels of dishonesty varied significantly across countries.*

A chi-squared test of proportion reporting heads against nationality rejects equality of proportions ($p < 0.001$). A Kruskal-Wallis test of mean quiz score against nationality rejects equality of medians ($p < 0.001$). Thus, both measures of dishonesty vary among nations.[12] Table 1 shows the pairs of countries whose proportions reporting heads differ in a chi-squared at 5% significance (corrected for false discovery rate), and table 2 shows the same for quiz scores. Overall differences are substantial: estimated dishonesty in the coinflip varies from 3.4% in Great Britain to 70% in China.

	RU	BR	IN, KR, JP	CN
GB	*	*	*	*
ZA		*	*	*
PT			*	*
GR, CH, DK, TR, US, AR				*

Table 1: Proportion reporting heads. Stars indicate differences at $p < 0.05$ corrected for false discovery rate

	GR, AR, RU	BR	PT, IN	CN	TR
JP	*	*	*	*	*
GB		*	*	*	*
US, DK			*	*	*
ZA, KR				*	*
GR, AR, RU, BR, PT					*

Table 2: Mean quiz scores. Stars indicate differences at $p < 0.05$ corrected for false discovery rate

To confirm these differences, I run regressions with controls. Tables 8 and 9 in the appendix show linear regressions of nationality dummies (with omitted category the US) on the probability of answering heads and on quiz score. The second column in each table includes as controls gender, age group, religious attendance, population density at the respondent's estimated location, incentive size (*Hilow*) and treatment order (*QuizFirst*). The third column includes further demographic controls, recorded in the second wave only: income, marital status, children, educational level, and a question on the importance of religion. Effects of nationality are robust. Few other demographic variables are significant. Those who report attending religious services more often are *more* likely to report heads, and score more in the quiz, but this effect is not present within wave 2. Older people scored less in the quiz but were not less likely to report heads. The incentive size treatment had

[12] This was also true among the eight wave 1 countries: heads against nationality, chi-squared test $p = 0.003$; Quiz score against nationality, Kruskal-Wallis test $p < 0.001$.

no effect. There is a small order effect in the quiz, with those who answered the quiz first scoring significantly more.

Country correlates of honesty

Now let us consider the link between country-level averages of dishonest behaviour and macro-level variables of interest. With just 15 observations, the analysis can only be correlational, leaving causal identification for future work. Another problem is that at country level, the correlation between mean quiz score and proportion reporting heads is positive but not significant (correlation 0.247, $p = 0.375$), despite the significant individual-level association. This is probably due to the small number of countries. Sometimes the two measures suggest different results: for example, in the coin flip the four least honest countries are the four Asian countries, China, India, South Korea and Japan, but Asian countries are not significantly more dishonest than others in the quiz, where Japan has the *lowest* mean score.[13] The Asia-other difference in the coin flip may be caused by cultural attitudes specific to this paradigm (for example, attitudes to gambling) rather than by differences in honesty *per se*. Paradigm-specific reactions like this are troubling for cross-national experiments: ideally, we would like multiple measures of the honesty construct with strong convergent validity. For now, rather than combining the measures, I report correlations with each of them separately, and place more confidence in correlations which are robust across the two.

Result 3 *Measures of dishonesty are negatively correlated with 2011 GDP per capita. This correlation is mainly driven by GDP growth prior to 1950.*

[13] In the first wave China and Japan were the least honest countries. India and South Korea were included in wave 2 to test the association with new data. The probability of India and South Korea having the highest scores out of the seven wave 2 countries, under the null, is $p = 2! \times 5!/7! = 0.048$.

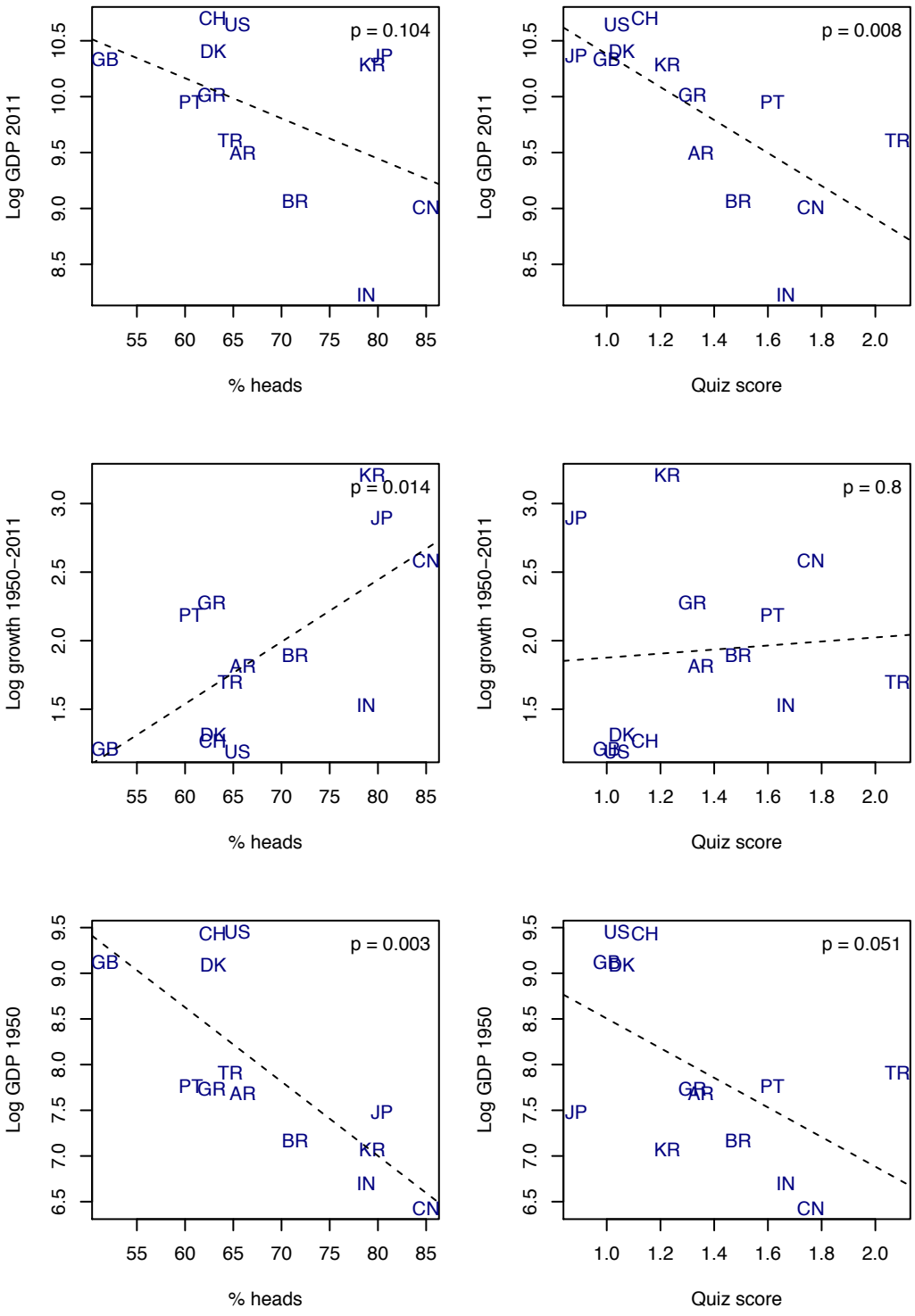

Figure 2: Dishonesty and GDP per capita. P values from bivariate linear regressions

Figure 2 plots country averages of each measure of dishonesty against three variables: GDP in 2011, GDP growth from 1950 to 2011 and GDP in 1950.[14] Lines from linear regressions are shown. Three facts stand out. First, there is a negative correlation between dishonesty and GDP in 2011, which is significant for the quiz but not the coin flip. Second, GDP growth from 1950 to 2011 shows either no correlation, or a positive correlation, with dishonesty. Lastly, the negative correlation exists and is significant for 1950 GDP. Table 3 shows linear regressions. Columns 1-2 and 4-5 show the bivariate relationships for 1950 and 2011 GDP. Columns 3 and 6 add 1950 GDP as a control to the 2011 GDP regression. Controlling for 1950 GDP, the coefficient on per cent reporting heads changes sign, while that for quiz scores remains negative and weakly significant, but shrinks substantially.

	Log GDP 1950	Log GDP 2011	Log GDP 2011	Log GDP 1950	Log GDP 2011	Log GDP 2011
Intercept	$13.50\ (1.45)^{***}$	$12.33\ (1.41)^{***}$	$2.95\ (3.08)$	$10.13\ (1.03)^{***}$	$11.85\ (0.63)^{***}$	$7.89\ (1.59)^{***}$
% heads	$-0.08\ (0.02)^{**}$	$-0.04\ (0.02)$	$0.02\ (0.02)$			
Quiz score				$-1.62\ (0.74)^{+}$	$-1.47\ (0.45)^{**}$	$-0.84\ (0.44)^{+}$
Log GDP 1950			$0.69\ (0.22)^{**}$			$0.39\ (0.15)^{*}$
R^2	0.58	0.22	0.62	0.30	0.49	0.70
Adj. R^2	0.54	0.15	0.54	0.24	0.45	0.64
Num. obs.	13	13	13	13	13	13
RMSE	0.70	0.68	0.50	0.90	0.55	0.44

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

Table 3: GDP regressions

Thus, the relationship between honesty and economic growth held over some time interval before 1950, but has been weaker or absent over the past sixty years. One story that fits this data is as follows: when institutions and technology are undeveloped, honesty is important as a substitute for formal contract enforcement. Countries that develop cultures putting a high value on honesty are able to reap economic gains. Later, this economic growth itself improves institutions and technology, making contracts easier to monitor and enforce, so that a culture of honesty is no longer necessary for further growth. However, since culture is highly persistent, the correlation between GDP and honesty remains visible in present-day behavioural data. Naturally, other interpretations are also compatible with the data, for example that the GDP-honesty correlation in 1950 was driven by an unobserved third variable. What can be said is that there is little evidence for a link between economic growth and honesty within the past 60 years.

One very simple explanation for this result would be that respondents in poor countries face a greater temptation to lie because they need the money reward more. Two pieces of evidence bear against this interpretation. First, the 66% increase from a \$3 to a \$5 incentive from the LOW to the HIGH treatment did not increase lying,

[14] GDP data is from Gleditsch (2002). Russia and South Africa are excluded as they do not have data in 1950.

either in the coin flip (heads reports: 69.4% in LOW, 66.9% in HIGH) or in the quiz (mean scores: 1.39 in LOW, 1.32 in HIGH).[15] Second, respondents in wave 2 reported their monthly income, and this also did not correlate with more lying (variable *Income* in tables 8 and 9).

I next examine possible roots of country-level dishonesty . One variable that might affect moral behaviour is religion, since different religions promulgate different norms. The classic study of Weber (2010) argues that Protestant religious teachings encouraged honesty in business dealings. Barro and McCleary (2003) argue that religion affects "traits such as honesty, thrift, willingness to work hard, and openness to strangers." Similarly, Porta et al. (1997) show that country-level measures of trust, from the World Values Survey questionnaire, are related negatively to membership in hierarchical religions: Catholicism, Orthodoxy and Islam.

Result 4 *Measures of dishonesty are negatively correlated with Protestantism at country level*

The percentage of Protestants in each country (Pew Research Center 2015) is correlated negatively and significantly with both measures of dishonesty (coin flip $\rho = -0.581$, $p = 0.023$; quiz $\rho = -0.571$, $p = 0.026$). Correlations with per cent Catholic are insignificant (coin flip $\rho = -0.23$, $p = 0.41$; quiz $\rho = 0.069$ $p = 0.808$).[16] Correlations with Protestantism become insignificant when controlling for 2011 GDP (coin flip $p = 0.354$, quiz $p = 0.266$). This is compatible with the assumption that Protestantism affects honesty, which in turn affects GDP.

Note that at individual level, there is no evidence that religious adherence is associated with honesty. Indeed, those who claim to attend services regularly, and that religion is very important to them, report heads *more* often and score *higher* in the quiz (appendix tables 8 and 9). Thus, countries with a Protestant cultural background have higher present-day honesty, but this difference is not necessarily reflected among present-day individuals.

Beliefs

Result 5 *Beliefs about countries' honesty were uncorrelated with their actual honesty*

Each respondent was asked how many respondents out of 100 would report heads in the coin flip, from one target country chosen randomly from all countries in the respondent's wave. Those in the INCENT condition were offered $2 if they guessed within 10 of the true number.

[15] This null result also held within individual countries. See Hugh-Jones (2015).

[16] There is not enough variation in the sample of countries to look at Orthodox Christianity.

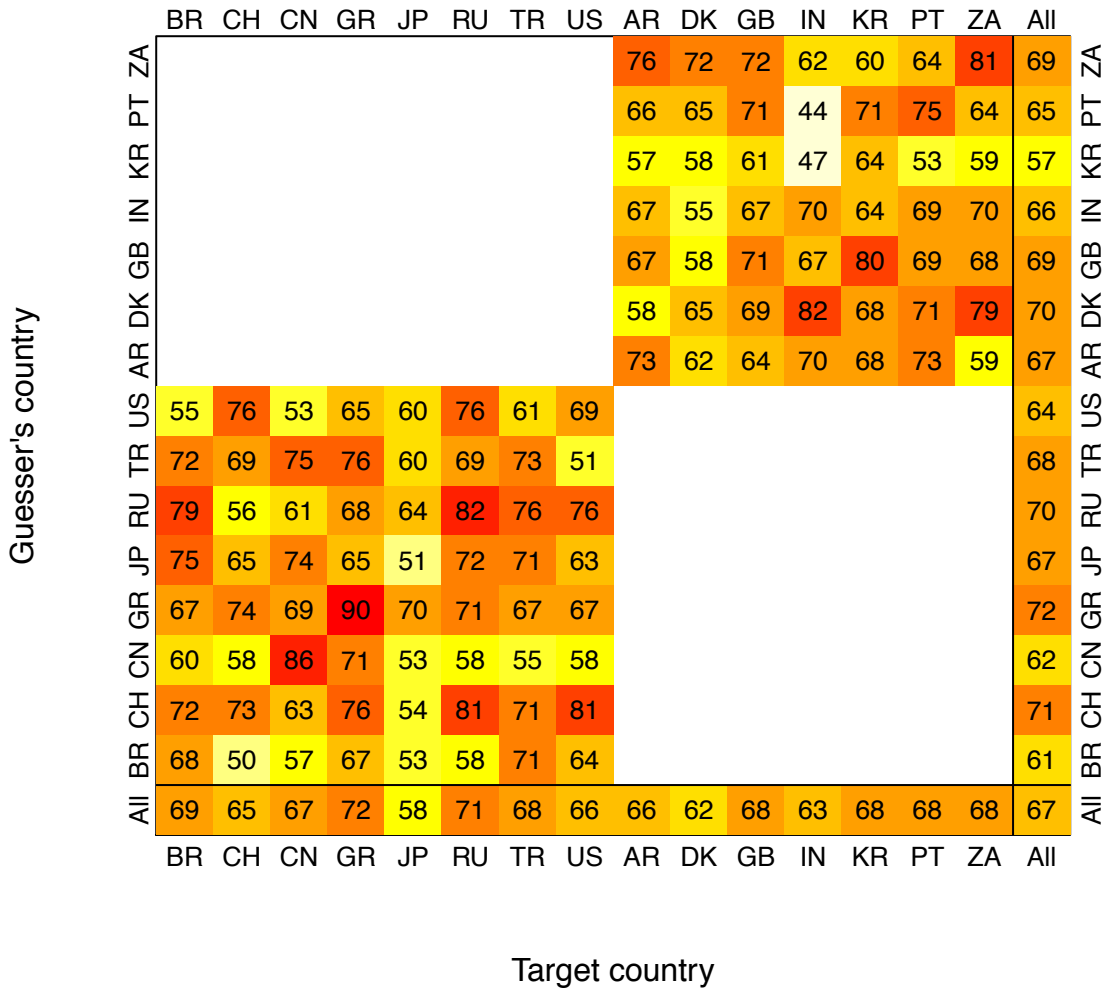

Figure 3: Average predictions of proportion reporting heads

Figure 3 shows average predictions about each country, by subjects from each country. Each cell is the average of guesses from only about 13 subjects, so standard errors are large. Nevertheless, some patterns in the data are already visible. Predictions vary systematically, both by target country (columns), and guesser's country (rows). And predictions are often higher along the diagonal, when subjects made guesses about their own country.

I next examine whether subjects' predictions reflect reality. Table 4 reports regressions of predicted number of heads on *TargetHeads*, the true percentage reporting heads in the target country. Column 1 shows the bivariate

correlation. The correlation between guesses and reality has the wrong sign. Column 2 repeats the regression for incentivized guesses only, with the same results. Clearly, expectations do not relate to reality. Indeed, at individual level, less than one in four subjects guessed within 10 percentage points of the true figure, in both incentivized and unincentivized treatments.

	All	Incentivized	All	Incentivized	All
(Intercept)	$74.91\ (4.79)^{***}$	$75.85\ (6.84)^{***}$	$88.16\ (6.33)^{***}$	$90.34\ (9.01)^{***}$	$88.02\ (6.39)^{***}$
TargetHeads	$-0.13\ (0.07)^{+}$	$-0.16\ (0.10)$	$-0.11\ (0.07)$	$-0.17\ (0.10)^{+}$	$-0.11\ (0.07)$
Heads			$4.10\ (1.46)^{**}$	$2.49\ (2.07)$	$4.19\ (1.55)^{**}$
QuizScore			$2.07\ (0.51)^{***}$	$2.18\ (0.73)^{**}$	$2.08\ (0.51)^{***}$
HomeHeads			$-0.28\ (0.07)^{***}$	$-0.28\ (0.10)^{**}$	$-0.28\ (0.07)^{***}$
TargetIsHome			$5.51\ (2.31)^{*}$	$9.58\ (3.25)^{**}$	$5.99\ (3.75)$
TargetDistance			$-0.23\ (0.17)$	$-0.07\ (0.24)$	$-0.23\ (0.17)$
Heads:TargetIsHome					$-0.69\ (4.21)$
R^2	0.00	0.00	0.04	0.05	0.04
Adj. R^2	0.00	0.00	0.04	0.04	0.04
Num. obs.	1379	689	1379	689	1379
RMSE	24.92	24.90	24.47	24.45	24.48

S.e.s clustered by target country and home country. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

Table 4: Coin flip: expectations vs. reality

What then drives expectations? The simplest story is that subjects may predict honesty on the basis of experience in their own country. Alternatively, subjects may "project" their own preferences and behaviours on to others: if so, those reporting heads or scoring high in the quiz may expect more dishonesty from others. Relatedly, Dieckmann et al. (2015) suggest that this "social projection" is more pronounced for social groups that are similar to one's own. Lastly, subjects' level of knowledge of other countries may affect their expectations, for example via reports of dishonesty or scandals in the media.

To test these theories, I added the following variables to the belief regression: *Heads*, whether the subject reported heads; subject's own *QuizScore*; *HomeHeads*, the proportion reporting heads in the subject's own country; *TargetIsHome*, which takes the value 1 if the target country was the subject's own and 0 otherwise; and *TargetDistance*, the distance between subject's country and the target country in thousands of kilometers, from the CEPII dataset (variable *distcap*; Mayer and Zignago 2011). *TargetDistance* is a proxy for subjects' level of knowledge of the target country, and for their judgments of similarity with their own country. The experience theory predicts a positive coefficient for *HomeHeads*. The projection theory predicts a positive coefficient for *Heads* and *QuizScore*, and possibly an interaction between *Heads* and *TargetIsHome*. Social identity theory (Tajfel et al., 1971) claims that humans attribute positive characteristics to their in-group; this would predict a negative coefficient on *TargetIsHome*. Broadly, if knowledge of other countries is relevant to judgments of

honesty, coefficients on *TargetIsHome* and *TargetDistance* should be significant.

Result 6 *Dishonest subjects believed others would be less honest on average. Subjects in dishonest countries believed others would be more honest. Subjects expected less honesty in their own country.*

Column 3 of table 4 shows the results. The *Heads* and *QuizScore* variables have the expected sign. Subjects who reported heads or scored highly in the quiz expected others to be more likely to report heads.[17] *Home-Heads*, however, has a significant negative sign, contradicting the experience theory: subjects from less honest countries took a more positive view of countries' honesty. *TargetIsHome* is positive and significant: on average, subjects expected their own countries to be less honest. Lastly, the coefficient on *TargetDistance* is negative but not significant. Column 4 repeats the exercise for only incentivized subjects, with essentially unchanged results. Thus, respondents showed three biases. Less honest subjects expected others to be less honest (social projection); subjects expected more honesty in other countries than their own; and subjects from more honest countries expected others to be less honest.

The positive coefficient on *TargetIsHome* runs counter to social identity theory. Dieckmann et al. (2015) suggests an explanation for a similar finding: subjects "socially project" their own behaviour on to fellow citizens more than on to foreigners, and since most subjects report heads, they expect others like them to do the same. To test this, column 5 repeats the regression, adding an interaction between *Heads* and *TargetIsHome*. The interaction is not significant and is negatively signed: subjects who reported heads did not believe their home country to be (relatively) less honest. So social projection seems not to be the explanation. One possibility, which cannot be tested within this data, is that subjects are more exposed to news stories about dishonesty in their own country than in other countries. The last result, that subjects from more honest countries are more pessimistic, is also puzzling. It is not solely driven by inter-country "snobbery": it holds even among subjects who guessed about their own countries (results not shown).

Conclusion

Comparing levels of honesty across 15 countries appears to confirm Feng Guifen's intuition: the connection between word and deed is less necessary in some places than in others. There is also some support for Feng's second claim, that this connection is good for economic development. This connection is clearer for earlier economic development, as summarized in 1950 GDP levels, than for growth in the past sixty years, consistently

[17] An alternative explanation is that subjects who really flipped heads naïvely expected others to do so too. But this does not explain the coefficient on *QuizScore*.

with the theory that honesty, and social capital more broadly, is a substitute for technologies and institutions that ease monitoring and accountability. Lastly, there is suggestive evidence of a link with cultural background in the form of religion.

In general it is an open question whether national stereotypes are accurate (Jussim et al., 2009). In this data, beliefs about other countries' honesty had little relation to the truth, or even to the average honesty of subjects' own countries. Instead, they appear to be driven by self-projection and other cognitive biases.

These results should be interpreted with caution. Respondents were not sampled randomly from national populations, and the two experimental measures of honesty were only weakly correlated at country level. These issues are shared with much cross-cultural experimental research. Experiments on representative samples in multiple countries are still relatively expensive and rare, and experimental economists have probably paid less attention than psychologists to questions of construct validity. To go beyond our current tentative conclusions, better sampling designs and measures will be necessary.

Despite these caveats, the large cross-country differences reported here seem unlikely to be explained by sample selection alone. There is a strong case that economic development has cultural roots (Banfield, 1958; Platteau, 2000; Guiso, Sapienza and Zingales, 2006). Experiments can help to isolate cultural differences, holding incentives and institutions constant. An important next step is to examine the historical process in more detail, so as to understand when and how culture matters, or mattered, to economic growth.

Acknowledgements

This work was supported by British Academy small grant SG122437. Tables were created using the *texreg* R package (Leifeld, 2013).

References

Abeler, Johannes, Anke Becker and Armin Falk. 2014. "Representative evidence on lying costs." *Journal of Public Economics* 113:96–104.

URL: *http://www.sciencedirect.com/science/article/pii/S0047272714000061*

Acemoglu, Daron and James Robinson. 2013. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Reprint edition ed. Crown Business.

Banfield, Edward C. 1958. *The moral basis of a backward society,*. Glencoe, Ill.; [Chicago: Free Press; Research Center in Economic Development and Cultural Change, University of Chicago.

Barro, Robert J. and Rachel McCleary. 2003. Religion and economic growth. Technical report National Bureau of Economic Research.

Berg, J., J. Dickhaut and K. McCabe. 1995. "Trust, reciprocity, and social history." *Games and Economic Behavior* 10(1):122–142.

Bornhorst, Fabian, Andrea Ichino, Oliver Kirchkamp, Karl H. Schlag and Eyal Winter. 2010. "Similarities and differences when building trust: the role of cultures." *Experimental Economics* 13(3):260–283.

Bucciol, Alessandro and Marco Piovesan. 2011. "Luck or cheating? A field experiment on honesty with children." *Journal of Economic Psychology* 32(1):73–78.
URL: *http://www.sciencedirect.com/science/article/pii/S0167487010001352*

Buchan, Nancy R., Rachel TA Croson and Robyn M. Dawes. 2002. "Swift Neighbors and Persistent Strangers: A Cross-Cultural Investigation of Trust and Reciprocity in Social Exchange1." *American Journal of Sociology* 108(1):168–206.

Charness, Gary and Martin Dufwenberg. 2006. "Promises and partnership." *Econometrica* 74(6):1579–1601.

Cowen, Tyler and Alexander Tabarrok. 2015. "The End of Asymmetric Information." *Cato Unbound* April 2015.
URL: *http://www.cato-unbound.org/2015/04/06/alex-tabarrok-tyler-cowen/end-asymmetric-information*

Croson, Rachel and Nancy Buchan. 1999. "Gender and culture: International experimental evidence from trust games." *American Economic Review* pp. 386–391.

Dieckmann, Anja, Urs Fischbacher, Veronika Grimm, Matthias Unfried, Verena Utikal and Lorenzo Valmasoni. 2015. Trust and beliefs among Europeans: Cross-country evidence on perceptions and behavior. Technical report IWQW Discussion Paper Series.

Dreber, Anna and Magnus Johannesson. 2008. "Gender differences in deception." *Economics Letters* 99(1):197–199.

Erat, Sanjiv and Uri Gneezy. 2012. "White lies." *Management Science* 58(4):723–733.

Fischbacher, Urs and Franziska Föllmi-Heusi. 2013. "Lies in disguise—an experimental study on cheating." *Journal of the European Economic Association* 11(3):525–547.

Gleditsch, K. S. 2002. "Expanded trade and GDP data." *Journal of Conflict Resolution* 46(5):712.

Gneezy, Uri. 2005. "Deception: The role of consequences." *American Economic Review* pp. 384–394.

Guiso, Luigi, Paola Sapienza and Luigi Zingales. 2006. "Does Culture Affect Economic Outcomes?" *Journal of Economic Perspectives* 20(2):23–48.
URL: *http://www.aeaweb.org/articles.php?doi=10.1257/jep.20.2.23*

Henrich, J. P. 2004. *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-scale Societies*. Oxford University Press.

Henrich, Joseph, Richard McElreath, Abigail Barr, Jean Ensminger, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins Gwako and Natalie Henrich. 2006. "Costly punishment across human societies." *Science* 312(5781):1767–1770.

Herrmann, Benedikt, Christian Thöni and Simon Gächter. 2008. "Antisocial punishment across societies." *Science* 319(5868):1362–1367.

Holm, Håkan J. and Anders Danielson. 2005. "Tropic trust versus nordic trust: Experimental evidence from tanzania and sweden*." *The Economic Journal* 115(503):505–532.

Holm, Håkan J. and Toshiji Kawagoe. 2010. "Face-to-face lying – An experimental study in Sweden and Japan." *Journal of Economic Psychology* 31(3):310–321.
URL: *http://www.sciencedirect.com/science/article/pii/S0167487010000085*

Hugh-Jones, David. 2015. "Ways to measure honesty: a new experiment and two questionnaires.".

Johnson, N.D. and A.A. Mislin. 2011. "Trust games: A meta-analysis." *Journal of Economic Psychology* 32(5):865–889.

Jussim, Lee, Thomas R. Cain, Jarret T. Crawford, Kent Harber and Florette Cohen. 2009. "The unbearable accuracy of stereotypes." *Handbook of prejudice, stereotyping, and discrimination* pp. 199–227.

Knack, Stephen and Philip Keefer. 1997. "Does social capital have an economic payoff? A cross-country investigation." *The Quarterly journal of economics* pp. 1251–1288.

Leifeld, Philip. 2013. "texreg: Conversion of Statistical Model Output in R to LaTeX and HTML Tables." *Journal of Statistical Software* 55(8):1–24.

URL: *http://www.jstatsoft.org/v55/i08/*

Mann, Heather, Ximena Garcia-Rada, Lars Hornuf, Juan Tafurt and Dan Ariely. 2015. "Cut from the Same Cloth: Surprisingly Honest Individuals Across Cultures.".

Mayer, Thierry and Soledad Zignago. 2011. "Notes on CEPII's distances measures: The GeoDist database.".

McCloskey, Deirdre N., Donald N. McCloskey and George K. Hersh Jr. 1990. *A Bibliography of Historical Economics to 1980*. Cambridge University Press.

McGarty, Craig, Vincent Y. Yzerbyt and Russell Spears. 2002. *Stereotypes as explanations: The formation of meaningful beliefs about social groups*. Cambridge University Press.

Pascual-Ezama, David, Toke R. Fosgaard, Juan Camilo Cardenas, Praveen Kujal, Robert Veszteg, Beatriz Gil-Gómez de Liaño, Brian Gunia, Doris Weichselbaumer, Katharina Hilken and Armenak Antinyan. 2015. "Context-dependent cheating: Experimental evidence from 16 countries." *Journal of Economic Behavior & Organization* .

Pew Research Center. 2015. "Christian Population as Percentages of Total Population by Country.".

URL: *http://www.pewforum.org/2011/12/19/table-christian-population-as-percentages-of-total-population-by-country/*

Platteau, Jean-Philippe. 2000. *Institutions, Social Norms, and Economic Development*. Psychology Press.

Porta, R. L, F. Lopez-de Silanes, A. Shleifer and R. W Vishny. 1997. "Trust in Large Organizations." *The American Economic Review* 87(2):333–338.

Soll, Jacob. 2015. "Germany's Destructive Anger." *The New York Times* .

URL: *http://www.nytimes.com/2015/07/15/opinion/germanys-destructive-anger.html*

Spence, Johnathan D. 2001. *The Search for Modern China*. New edition ed. W. W. Norton & Co.

Tajfel, H., M. G Billig, R. P Bundy and C. Flament. 1971. "Social categorization and intergroup behavior." *European Journal of Social Psychology* 1(2):149–178.

The Economist. 2014. "The Big Mac index.".

URL: *http://www.economist.com/content/big-mac-index*

Vanberg, Christoph. 2008. "Why do people keep their promises? An experimental test of two explanations1." *Econometrica* 76(6):1467–1480.

Weber, Max. 2010. *The Protestant Ethic and the Spirit of Capitalism*. CreateSpace Independent Publishing Platform.

Whiteley, Paul. 2012. "Are Britons Getting More Dishonest?".
URL: *http://www.west-info.eu/files/Research11.pdf*

Yamagishi, Toshio. 1988. "The provision of a sanctioning system in the United States and Japan." *Social Psychology Quarterly* pp. 265–271.

Appendix A: descriptive statistics

Variable	Levels	n	%	∑%
Nationality	AR	97	7.0	7.0
	BR	98	7.0	14.0
	CH	70	5.0	19.1
	CN	100	7.2	26.3
	DK	89	6.4	32.7
	GB	89	6.4	39.1
	GR	94	6.8	45.8
	IN	99	7.1	52.9
	JP	97	7.0	59.9
	KR	102	7.3	67.3
	PT	99	7.1	74.4
	RU	100	7.2	81.6
	TR	85	6.1	87.7
	US	81	5.8	93.5
	ZA	90	6.5	100.0
	all	1390	100.0	
Age	18-24	274	19.7	19.7
	25-34	287	20.6	40.4
	35-44	281	20.2	60.6
	45-54	289	20.8	81.4
	55+	259	18.6	100.0
	all	1390	100.0	
Gender	Female	705	50.8	50.8
	Male	682	49.2	100.0
	all	1387	100.0	
ReligAttend	Never	544	39.2	39.2
	Sometimes	637	45.9	85.1
	Weekly	207	14.9	100.0
	all	1388	100.0	
Heads	0	443	31.9	31.9
	1	947	68.1	100.0
	all	1390	100.0	
QuizScore	0	588	42.3	42.3
	1	156	11.2	53.5
	2	209	15.0	68.6
	3	437	31.4	100.0
	all	1390	100.0	

Table 5: Descriptive statistics

Variable	**Levels**	**n**	%	∑%
MaritalStatus	Single	283	42.6	42.6
	Married	312	47.0	89.6
	Sep.	12	1.8	91.4
	Div.	43	6.5	97.9
	Wid.	14	2.1	100.0
	all	664	100.0	
HasChildren	0	271	40.8	40.8
	1	393	59.2	100.0
	all	664	100.0	
AgeLeftEduc	16	31	4.7	4.7
	17	36	5.4	10.1
	18	73	11.0	21.1
	19	44	6.6	27.7
	20	48	7.2	34.9
	21+	334	50.2	85.1
	Still in education	90	13.5	98.6
	Under 16	9	1.4	100.0
	all	665	100.0	
Trust	Most people can be trusted	218	32.9	32.9
	You need to be very careful	444	67.1	100.0
	all	662	100.0	
ReligImportance	No religion	133	20.1	20.1
	Not important	173	26.1	46.1
	Quite important	208	31.4	77.5
	Very important	149	22.5	100.0
	all	663	100.0	

Table 6: Descriptive statistics (wave 2 demographics)

Variable	Min	Median	Mean	Max
Integrity	15.0	53.0	51.8	60.0
Guess	0.0	70.0	66.4	100.0
Income	0.0	1.3	2.0	31.4

Table 7: Descriptive statistics (continuous variables)

Appendix B: regressions

	Pct heads	Pct heads	Pct heads
(Intercept)	0.52 (0.05)***	0.49 (0.06)***	0.28 (0.17)
NationalityAR	0.14 (0.07)*	0.11 (0.07)$^{+}$	0.10 (0.11)
NationalityBR	0.20 (0.07)**	0.18 (0.07)**	
NationalityCH	0.11 (0.07)	0.10 (0.07)	
NationalityCN	0.33 (0.07)***	0.33 (0.07)***	
NationalityDK	0.11 (0.07)	0.10 (0.07)	0.10 (0.10)
NationalityGR	0.11 (0.07)	0.08 (0.07)	
NationalityIN	0.27 (0.07)***	0.22 (0.08)**	0.40 (0.12)**
NationalityJP	0.29 (0.07)***	0.29 (0.07)***	
NationalityKR	0.28 (0.07)***	0.26 (0.07)***	0.23 (0.10)*
NationalityPT	0.09 (0.07)	0.08 (0.07)	0.06 (0.11)
NationalityRU	0.19 (0.07)**	0.17 (0.07)*	
NationalityTR	0.13 (0.07)$^{+}$	0.09 (0.07)	
NationalityUS	0.14 (0.07)$^{+}$	0.12 (0.07)$^{+}$	
NationalityZA	0.01 (0.07)	−0.03 (0.07)	0.05 (0.11)
GenderMale		0.01 (0.02)	0.04 (0.05)
Age25-34		0.02 (0.04)	0.12 (0.09)
Age35-44		−0.03 (0.04)	−0.11 (0.10)
Age45-54		−0.01 (0.04)	−0.08 (0.10)
Age55+		−0.00 (0.04)	−0.04 (0.11)
ReligAttendSometimes		0.06 (0.03)*	0.13 (0.07)$^{+}$
ReligAttendWeekly		0.12 (0.04)**	0.15 (0.11)
PopDensity		−0.00 (0.00)	−0.00 (0.00)
QuizFirst		−0.03 (0.02)	0.04 (0.05)
HilowLOW		0.03 (0.02)	0.07 (0.05)
Income			0.01 (0.01)
HasChildren			0.05 (0.07)
MaritalStatusMarried			−0.08 (0.07)
MaritalStatusSep.			0.22 (0.18)
MaritalStatusDiv.			−0.05 (0.11)
MaritalStatusWid.			0.15 (0.18)
ReligImportanceNot important			0.01 (0.09)
ReligImportanceQuite important			0.09 (0.10)
ReligImportanceVery important			0.26 (0.12)*
Education<21			0.01 (0.06)
EducationStudentNow			−0.11 (0.09)
R^2	0.04	0.05	0.13
Adj. R^2	0.03	0.04	0.06
Num. obs.	1390	1385	383
RMSE	0.46	0.46	0.46

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

Table 8: Coin Flip Regressions
Omitted categories: Nationality, GB; Age, 18-24; ReligAttend, never; Education, finished education at 21+; MaritalStatus, single; ReligImportance, no religion. PopDensity is population density at respondent's location (estimated from IP address).

	Quiz score	Quiz score	Quiz score
(Intercept)	1.00 (0.14)***	0.90 (0.16)***	1.19 (0.46)*
NationalityAR	0.35 (0.19)$^{+}$	0.20 (0.19)	0.17 (0.30)
NationalityBR	0.49 (0.19)**	0.40 (0.18)*	
NationalityCH	0.14 (0.20)	0.10 (0.20)	
NationalityCN	0.76 (0.19)***	0.76 (0.18)***	
NationalityDK	0.06 (0.19)	0.02 (0.19)	−0.13 (0.28)
NationalityGR	0.32 (0.19)$^{+}$	0.14 (0.20)	
NationalityIN	0.67 (0.19)***	0.43 (0.21)*	0.82 (0.33)*
NationalityJP	−0.11 (0.19)	−0.15 (0.19)	
NationalityKR	0.23 (0.18)	0.15 (0.18)	0.32 (0.27)
NationalityPT	0.62 (0.19)***	0.51 (0.19)**	0.46 (0.29)
NationalityRU	0.41 (0.19)*	0.29 (0.19)	
NationalityTR	1.08 (0.19)***	0.88 (0.19)***	
NationalityUS	0.04 (0.20)	−0.08 (0.19)	
NationalityZA	0.19 (0.19)	0.03 (0.19)	−0.03 (0.29)
GenderMale		0.03 (0.07)	−0.10 (0.14)
Age25-34		0.13 (0.11)	0.22 (0.25)
Age35-44		−0.15 (0.11)	−0.27 (0.27)
Age45-54		−0.25 (0.11)*	−0.55 (0.28)*
Age55+		−0.46 (0.11)***	−0.71 (0.30)*
ReligAttendSometimes		0.26 (0.08)***	0.02 (0.20)
ReligAttendWeekly		0.35 (0.11)**	0.27 (0.30)
PopDensity		0.00 (0.00)	−0.00 (0.00)
QuizFirst		0.21 (0.07)**	0.07 (0.13)
HilowLOW		0.10 (0.07)	0.18 (0.13)
Income			0.01 (0.03)
HasChildren			0.05 (0.19)
MaritalStatusMarried			−0.22 (0.20)
MaritalStatusSep.			0.96 (0.49)*
MaritalStatusDiv.			−0.10 (0.30)
MaritalStatusWid.			−0.11 (0.48)
ReligImportanceNot important			−0.16 (0.24)
ReligImportanceQuite important			0.19 (0.28)
ReligImportanceVery important			0.04 (0.33)
Education<21			0.17 (0.15)
EducationStudentNow			−0.36 (0.26)
R^2	0.06	0.10	0.15
Adj. R^2	0.05	0.08	0.08
Num. obs.	1390	1385	383
RMSE	1.27	1.25	1.26

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.1$

Table 9: Quiz Regressions
Omitted categories: Nationality, GB; Age, 18-24; ReligAttend, never; Education, finished education at 21+; MaritalStatus, single; ReligImportance, no religion. PopDensity is population density at respondent's location (estimated from IP address).

Appendix C: further results

I report one further result here to mitigate the desk drawer problem.

Result 7 *Mean quiz scores correlate negatively but insignificantly with per-country trust from the World Values Survey*

The World Values Survey includes a question on social trust. Figure 4 plots the proportions responding "Most people can be trusted" against quiz scores, for the 12 countries for which data is available (World Values Survey waves 5-6). The slope is negative and insignificant ($\beta = -8.14$, $p = 0.576$). China is an outlier. Respondents in China, a non-democracy where freedom of expression is severely restricted, may not answer truthfully. The slope excluding China is negative and weakly significant ($\beta = -21.555$, $p = 0.065$). However, the correlation of trust with per cent reporting heads is not significant, with or without China ($p = 0.298$ and 0.897 respectively).

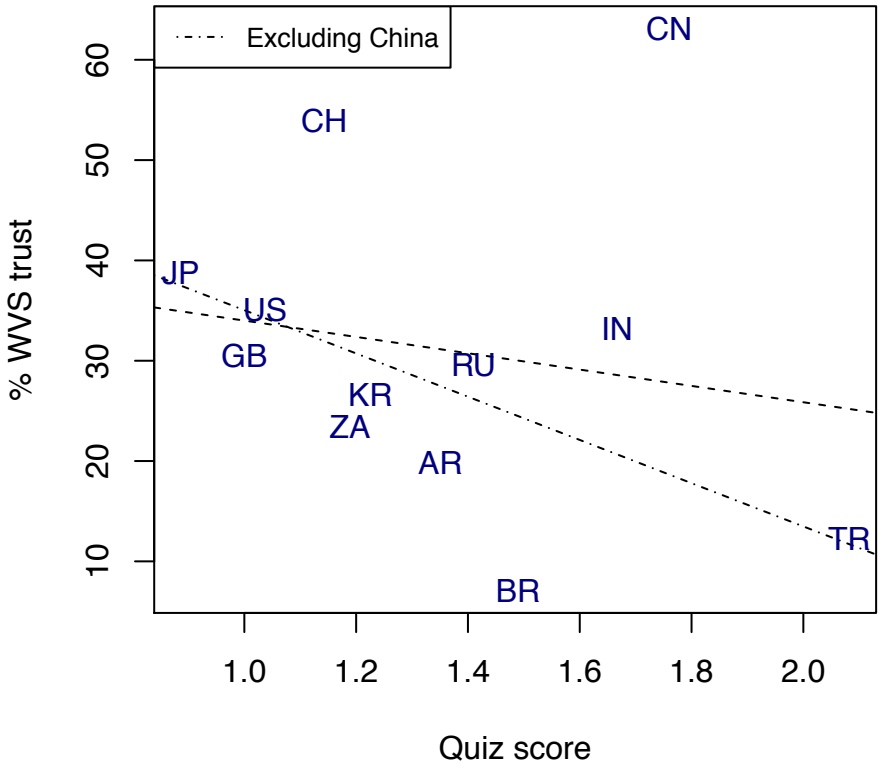

Figure 4: Quiz scores and trust

Appendix D: experiment instructions

In the next part of the survey, you will be asked to make some decisions. The decisions you make will affect the amount of money that you receive.

You do not have to fill in this part of the survey if you do not want to: in this case, simply leave responses blank. Data from this part of the survey will be stored at the University of Essex and may be released to other researchers. However, this data will not include your name or any other information that could identify you.

I consent to taking part in this part of the survey [ ]

What is your gender?

[ ] Male [ ] Female

How old are you?

[ ] Under 18 [ ] 18-24 [ ] 25-34 [ ] 35-44 [ ] 45-54 [ ] 55 or over

In which country do you reside? ...

Below is a short quiz on the topic of music. We will pay you $INCENTIVE extra if you get all the questions right. This payment will be added to your standard payment for taking the survey. Please answer the questions on your own, without looking them up elsewhere.

Who wrote the composition "Für Elise"?

What is Lady Gaga's real first name?___

Name the drummer of the rock group Nirvana.___

In what year was Claude Debussy born?___

How many valves are there on a standard modern trumpet?___

Name the town and state of the US where Michael Jackson was born.___

I confirm that I answered the quiz on my own, and that I did not look up any answers. [ ]

On the next page you will need to flip a coin.

Please get a coin ready before moving on!

OK, I have a coin ready [ ]

Please now flip the coin. If the coin lands on "heads", we will pay you an extra $INCENTIVE.

This payment will be added to your standard payment for taking the survey.

Did the coin land on heads?

[ ] Yes [ ] No

Please think about each of the following actions. In your opinion, is the action always justified, sometimes justified, rarely justified or never justified?

Always justified [ ] Sometimes justified [ ] Rarely justified [ ] Never justified [ ] Don't know [ ]

Claiming government benefits to which you are not entitled

Buying something which you know is stolen

Taking cannabis

Keeping money that you found in the street

Lying in your own interests

Having an affair when you are married

Having sex under the legal age of consent

Failing to report accidental damage you've done to a parked vehicle

Throwing away litter in a public place

Driving under the influence of alcohol

Avoiding a fare on public transport

Cheating on taxes if you have a chance

Someone accepting a bribe in the course of their duties

Driving faster than the speed limit

Making up things on a job application

Some people from RANDOM COUNTRY are answering this questionnaire.

Think back to the coin flip question. People from RANDOM COUNTRY were paid approximately the same amount as you if they flipped heads.

What percentage (%) of people from RANDOM COUNTRY do you think will report that the coin landed on heads?___

[If in INCENTIVIZED GUESS treatment] *We will pay you an extra 2$ if your answer is within 10% of the correct percentage.*

Please enter your nationality:___

Do you attend religious services?

[ ] Yes, at least weekly [ ] Yes, sometimes [ ] No, never

[Only for second wave countries:]

Please indicate your marital status:

[ ] Single [ ] Married [ ] Separated [ ] Divorced [ ] Widowed

How many children do you have (including step-children)?

[ ] 0 [ ] 1 [ ] 2 [ ] 3 [ ] 4 or more

What is your total monthly income (from all sources, net of tax)?___

Is religion an important part of your life?

[ ] Yes, very important [ ] Yes, quite important [ ] No, not important [ ] I have no religion

At what age did you complete your education?

[ ] Under 16 [ ] 16 [ ] 17 [ ] 18 [ ] 19 [ ] 20 [ ] 21 or over [ ] I have not yet completed my education

Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?

[ ] Most people can be trusted [ ] You need to be very careful

Which of these things, if any, have you done in the past 12 months?

[ ] Done in past year [ ] Not done in past year

Avoided a fare on public transport

Made something up on a job application

Downloaded music or videos without paying for them

Called in sick to work when not actually unwell