DISCUSSION PAPER SERIES

No. 7717

SIMILARITIES AND DIFFERENCES WHEN BUILDING TRUST: THE ROLE OF CULTURES

Fabian Bornhorst, Andrea Ichino,
Oliver Kirchkamp, Karl Schlag and
Eyal Winter

PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

Available online at: www.cepr.org/pubs/dps/DP7717.asp

ISSN 0265-8003

SIMILARITIES AND DIFFERENCES WHEN BUILDING TRUST: THE ROLE OF CULTURES

Fabian Bornhorst, European University Institute
Andrea Ichino, Università di Bologna, CESifo, IZA and CEPR
Oliver Kirchkamp, Universität Jena
Karl Schlag, Universitat Pompeu Fabra
Eyal Winter, Hebrew University of Jerusalem

Discussion Paper No. 7717
February 2010

Centre for Economic Policy Research
53–56 Gt Sutton St, London EC1V 0DG, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 7717

February 2010

ABSTRACT

Similarities and Differences when Building Trust: the Role of Cultures

We run an experiment in which students of different European nationalities are matched in groups of five and repeatedly choose with whom within their group they want to play a trust game. Participants observe of each other age, gender, nationality and number of siblings. The region of origin, "North" or "South" is a major determinant of success in the experiment. Participants tend to trust those they trusted before and who trusted them. We do not find evidence of regional discrimination per se. It is only the underlying and significant differences in behavior that translate through repeated interactions into differences in payoffs between the two regions.

JEL Classification: C91, C92 and Z13
Keywords: European regions, experiments, trust and trustworthiness

Fabian Bornhorst
Department of Economics
European University Institute
Villa San Paolo
Via della Piazzuola 43
I-50133 Florence (FI)
ITALY

Email: fabian.bornhorst@iue.it

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=160844

Andrea Ichino
Dipartimento di Scienze Economiche
Università di Bologna
Piazza Scaravilli 2
40126 Bologna
ITALY

Email: andrea.ichino@unibo.it

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=111104

Oliver Kirchkamp
Universität Jena
Wirtschaftswissenschaftliche Fakultät
Lehrstuhl für Empirische und
Experimentelle Wirtschaftsforschung
Carl-Zeiß-Straße 3
07743 Jena
GERMANY

Email: oliver@kirchkamp.de

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=152135

Karl Schlag
Universitat Pompeu Fabra
Department of Economics and
Business
Ramon Trias Fargas 25-27
08005 Barcelona
SPAIN

Email: karl.schlag@upf.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=135752

Eyal Winter
The Hebrew University of Jerusalem
Mount Scopus
Jerusalem 91905
ISRAEL

Email: mseyal@mscc.huji.ac.il

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=160845

Submitted 21 February 2010

We would like to thank the Research Council of EUI for financing this study, Francesco Caselli, Colin Crouch, Jaap Dronkers, Margaret Meyer, Massimo Motta, Soren Johansen, Jacques Ziller and seminar participants at Bocconi, the CEPR-PPS symposium, EUI, IIES, IZA, University of Bologna, University of Milan and UPF for insightful comments. We also thank Jessica Spataro for editorial assistance. Karl Schlag gratefully acknowledges financial support from the Department of Economics and Business of the Universitat Pompeu Fabra, Grant AL 12207, and from the Spanish Ministerio de Educacion y Ciencia, Grant MEC-SEJ2006-09993.

Similarities and Differences when Building Trust: the Role of Cultures*

Fabian Bornhorst‡ Andrea Ichino§ Oliver Kirchkamp¶
Karl H. Schlag‖ Eyal Winter**

17th February 2010

Abstract

We run an experiment in which students of different European nationalities are matched in groups of five and repeatedly choose with whom within their group they want to play a trust game. Participants observe of each other age, gender, nationality and number of siblings. The region of origin, "North" or "South" is a major determinant of success in the experiment. Participants tend to trust those they trusted before and who trusted them. We do not find evidence of regional discrimination per se. It is only the underlying and significant differences in behavior that translate through repeated interactions into differences in payoffs between the two regions.

JEL-Code: C91, C92, Z13
Keywords: Trust, trustworthiness, European regions, experiments

1 Introduction

Most economic interactions are preceded by a stage in which agents select partners. Entrepreneurs select their counterparts for a partnership, firms select suppliers, consumers

*We would like to thank the Research Council of EUI for financing this study, Francesco Caselli, Colin Crouch, Jaap Dronkers, Margaret Meyer, Massimo Motta, Soren Johansen, Jacques Ziller and seminar participants at Bocconi, the CEPR-PPS symposium, EUI, IIES, IZA, University of Bologna, University of Milan and UPF for insightful comments. We also thank Jessica Spataro for editorial assistance. Karl Schlag gratefully acknowledges financial support from the Department of Economics and Business of the Universitat Pompeu Fabra, Grant AL 12207, and from the Spanish Ministerio de Educacion y Ciencia, Grant MEC-SEJ2006-09993.

‡European University Institute.
§University of Bologna, CEPR, CESifo and IZA.
¶University of Jena.
‖Universitat Pompeu Fabra.
**The Hebrew University of Jerusalem.

select retailers and employers choose workers from pools of applicants. Trust and reciprocity play an important role in such interactions and the (initial) choice of a partner as well as the decision about the volume of activity to a large extent depend on the agent's beliefs about the prospects of building trust and reciprocity with potential partners. If the interaction takes place repeatedly, past experience will play a role as well. Agents are expected to return to those partners who proved to be trustworthy, and avoid those who failed to reciprocate.

In this paper we report on experimental results that emphasize the role of cultural diversity in shaping agents' choices of partners as well as in determining the outcomes of their economic interactions. We run an experiment in which participants are matched to a group of five and repeatedly have to choose with whom within their group they want to play a trust game. Within this group participants observe of each other age, gender, nationality and number of siblings. We find that difference in success in the experiment can only be explained in terms of national membership, specifically in terms of belonging to northern or southern Europe.

Our analysis adds to the growing literature suggesting that differences in "culture" may explain differences in economic performance across communities. With respect to research based on observational data, our experimental setting allows us to study, in a controlled environment, the "anatomy" and the "mechanics" of how culture affects performance through the selection of partners and through reciprocal behavior in a setting of repeated interactions. Thus, with the word "culture" we have in mind here the set of social norms and individual beliefs that sustain equilibria as focal points in repeated social interactions.[1] Our results show that in a global environment where economic interactions go across countries and cultures, national diversity may have a substantial impact on agents' initial beliefs regarding trust and trustworthiness of partners as well as on the evolution of their interaction over multiple transactions.

The importance of trust for economic success in a society has been emphasized by many authors, like for example Arrow (1972), Fukuyama (1995), Putnam (1993), Knack and Keefer (1997), La Porta et al. (1997), and Guiso et al. (2004a,b). We follow Glaeser et al. (2000) in using the "trust game experiment" designed by Berg et al. (1995) to measure trust and trustworthiness in a multi-cultural environment.

The standard format of a trust game involves two players. The "sender", who is assigned an amount of money x by the experimenter, decides on a transfer $0 \leq t \leq x$ to

[1]See Schotter (1981) and Greif (1994). From this viewpoint, our paper relates also to the recent debate on the role of history, culture and institutions as determinants of economic development. See, among others, Hall and Jones (1999), La Porta et al. (1999), Acemoglu, Johnson and Robinson (2001), Glaeser et al. (2004) and Tabellini (2005).

be made to the "receiver", who receives three times the amount of this transfer. Thus, if the sender concedes the amount t to the receiver, then the latter receives $3t$ (while the sender loses just t). Following the transfer made by the sender, the receiver decides how much she wants to return. The amount that the receiver decides to return is denoted by g, $g \in [0, 3t]$, and is equal to what the sender gets back. The unique Nash equilibrium outcome of the game is for the receiver to make zero payback and therefore for the sender to make no transfer at all. Berg et al. (1995) finds that senders make considerable transfers but that these were not backed by payback behavior. Among other papers that study behavior in this trust game is Buchan et al. (2006), which involves a comparison across different countries including the US, China, Japan and Korea. Their paper focuses on the effect of preliminary discussions within groups on behavior in the trust game. Holm and Danielson (2005) compare behavior in the trust game and the dictatorship game between Sweden and Tanzania. Behavior in the two countries is significantly different, particularly in the relationship between the way they played the trust game and the way they answered survey questions concerning attitudes towards trust.

Our framework differs from this strand of literature in three major aspects. Firstly, our main focus is on the dynamic of trust and cooperation. We design a dynamic version of the trust game that allows trust and reciprocity to build up and evolve. Secondly, we are interested in studying the heterogeneity among players in terms of trust, reciprocity and the level of being trusted. Finally, we wish to address the above issue in a fully multi-cultural set up, where players of different nationalities interact in the same strategic environment. In particular, in contrast to standard trust games in which the matching between players is fixed, we wish to highlight the role of the choice of a partner in real settings. We do so by allowing participants to choose the partner to whom they make a transfer. Each person in a group of five individuals acts as a sender and as a receiver. As a sender participants choose both the person to whom a transfer is made and the amount of transfer. If chosen as a receiver by one or more fellow players receivers choose how much to transfer back to each sender. These choices are made after receiving a list of characteristics on each of the group members, including nationality, gender, age and number of siblings.

Related to our framework are three papers which report results on a one shot trust game played between observable cultural groups; Fershtman and Gneezy (2000, Ashkenazi and Sepharadi), Willinger et al. (2003, French and German) and Fershtman et al. (2005, Flemish vs Walloons and Nonorthodox and Secular vs Ultraorthodox Jews).[2] In contrast

[2]Other examples are Holm and Nystedt (2005), Slonim (2006), Slonim and Garbarino (2008). However, these are based on age and gender and not on cultural traits.

to their framework in which matching is fixed and the interaction involved a one shot game, in our framework each participant can act both as a sender and as a receiver; participants choose their partner and interact repeatedly within the same group. These features allow us not only to detect unequal treatment but also to go more deeply into its roots by analyzing the way it evolves over different periods.[3]

Our pool of participants involves Ph.D. students at the European University Institute (EUI) in Florence. The EUI whose main objective is to provide advanced academic training to Ph.D. students in a European perspective, attracts young intellectuals from EU member countries with substantial international exposure and with a typical fluency in at least three European languages. If the role of nationality within this group is strong, we would expect it to be even stronger among the general population of Europe. Moreover we show that this role does not arise because of discrimination based on tastes, but is instead rooted in behavioral differences of players that correlate significantly with nationality. These differences would probably not produce major effects in a static environment. Their effects on outcomes emerge through the dynamic selection of partners and through repeated interactions. In other words, our results, presented in Section 3, indicate a major heterogeneity in players' behavior. And this heterogeneity translates into a significant North and South difference in terms of overall earnings in the game with North earning substantially more than South.

We propose three explanations for this finding: A) **Investment:** North are better investors as they transfer more and thus they are paid back more since players return roughly half of what they receive. B) **Networking:** Due to North showing a high level of trust (by transferring a lot) North enjoys more contacts and higher transfers in subsequent periods, either as a form of reciprocation or because players attribute more trust to North. C) **Selection:** North are doing better in selecting partners. Specifically, they show more bias towards those who revealed trust in the past, and gain from the fact that trust and trustworthiness are correlated.

We point out that Investment is a static effect as it relies on reciprocity in a single period. In contrast, Networking and Selection are dynamic effects. The networking effect concerns with the way players' current actions affect their future payoffs, while selection concerns the way past experience affect current actions. Our analysis will also show that the dynamic effects outweigh the static effect in generating the earning differences

[3]Along these lines Cochard et al. (2004), King-Casas et al. (2005), Schotter and Sopher (2006) and Charness et al. (2009) also consider repeated interaction but differ as they do not allow for endogenous partner selection. Related, Holm and Nystedt (2005) and Slonim and Garbarino (2008) ask participants to indicate preferred partners. In contrast, we explicitly allow them to select their partner and infer their preferences from these choices.

between North and South. To establish these result we shall go through some intermediate observations that will include the following:

1. Overall North shows a higher level of trust than South in all but the last period of the session (in which a strong end-effect prevails).

2. Trust breeds trust, i.e., players who show a high level of trust at early periods are more likely to be contacted later.

3. Both North and South favor more those that were previously trustworthy or who had previously trusted them. However, North is more sensitive than South to previous trust of others.

4. Rather than dying out, the earning gap between North and South builds up through the dynamics of the game. More specifically, North's earning is growing faster and more steadily across periods than that of South.

Note that the simultaneous presence of effects 1, 2, and 3 is an equilibrium phenomenon. In a society where people reciprocate to trust it pays off to trust, and in order to sustain and maintain it people have to reciprocate to trust. Likewise lack of trust jointly with the lack of reciprocity to trust forms an equilibrium as well. Hence, one can view cultural differences regarding trust as a prevalence of multiple equilibria.

We interpret our findings and particularly result 4 as an indication that cultural differences in standards regarding trust and reciprocity may be sufficiently robust to determine economic performance in a significant and persistent way. It is beyond the scope of our paper to identify the origin of these cultural differences but we propose the possibility that at least result 1 may emerge merely from an income effect inasmuch as "trust", "generosity" and "reciprocity" are luxury (normal) goods. It may also be related to the fact that risk aversion typically declines with income. A non-necessarily alternative interpretation, (and indeed one that can better explain result 3) suggested by Banfield (1958) and Putnam (1993) in a different context, is that these differences finds their roots in distant history and traditions, for example concerning the role of family and social capital in the two regions. In cultures where many of the social and economic interactions are conducted around the family, where trust is almost biologically built in, networking skills are not as essential as in cultures where most economic interactions are conducted outside the family.

The paper is organized as follows. In Section 2 the experimental design is introduced. Section 3 contains the analysis subdivided into population behavior, individual heterogeneity, differences in success across personal characteristics, a closer look at the success of North, differences in behavior across regions, observability of regions and dynamic choice. In Section 4 we conclude. The appendix to the paper contains more details on the experimental design and on the conditional logit model used to analyze learning.

2 The Design

The design of our experiment is described extensively in the appendix A. Here we limit ourselves to a summary of its most important features. We conducted 4 sessions with a total of 110 participants hired in 2002 among EUI Ph.D. students from different European countries. The number of participants in each session was a multiple of 5. Upon entry, participants were asked to complete a form on the computer screen in which they had to specify personal information that consisted of age, gender, nationality, and number of siblings. Despite the provision of this personal information, we took some measures to ensure anonymity. When the experimental protocol (see below) made personal information observable to others we added a small random mean preserving term to the true age and then only revealed the perturbed age. Participants could only take part in the experiment if sufficiently many other EUI students had the same characteristics. Moreover, participants were randomly spread across different rooms and buildings, while matchings were designed so to make sure that participants playing in the same group never had eye contact with any member of the group they were assigned to. We have used all the measures described above in order to confidently rule out the possibility that participants identify a person in their groups beyond the information revealed to them. Participants interviewed after the experiments confirmed that they had not been able to identify members in their group.

The actual experiment consists of four identical supergames which we refer to as repetitions. Each repetition consists of 6 periods. At the beginning of such a repetition participants are randomly matched into groups of five players and remain within this group throughout the repetition. The personal information about the other players is then made public within the group with the caveat that only the perturbed age is revealed. Play within a repetition consists of 6 periods. At the beginning of each period, each player receives an endowment of 100 tokens. Participants are then given the opportunity to choose one of the four other players from the same group and to transfer any part of the initial endowment to the player chosen. Choosing no player is also possible. In this case the transfer is equal to zero. If a sender makes a transfer of t to a receiver, the latter receives a transfer equal to $3t$. Each participant who receives a transfer has the option to return any part of it to the person who made the transfer. Then the period ends. All decisions are made via computer terminals. Figures 6 to 9 in the appendix provide examples of the screenshots seen by participants. At the end of each period a participant sees on the screen only the actions and payoffs of the interactions in which she was involved (i.e., the one in which she was a sender if she transferred a positive amount to some player, and the ones in which she was a receiver if she received a transfer from one or more other players). Thus, participants do not know at the end of a period what happened between other pairs of players.[4] At the end of each repetition each participant is also informed about the total profit she made in the 6 periods of that repetition. In the original experiment two further repetitions followed these four but this data is not used in this paper.

As in Berg et al. (1995), in reporting the results we refer to two characteristics of players: "trust" and "trustworthiness". Trust concerns sending behavior. It refers to the willingness of a player to contact another player and to make high transfers, which we interpret as a propensity to trust the receiver to reciprocate.[5] Trustworthiness stands for the tendency of a player to reciprocate by making a generous payback for a transfer she received. Trustworthiness will be measured as the average return ratio, i.e. what receivers return to senders as a fraction of what they have received. The precise formulae will be explained later.

We point out that the dynamic feature of the trust game that we design does not alter the equilibrium prediction of the game. As in the standard Berg et al. (1995) game also here no player interested only in own payoffs should make any positive transfer at any period of the game. The reason is that the number of periods in each repetition is finite and commonly known, hence one can start with the last period and invoke a standard backward induction argument. Of course this prediction need not carry over in richer models. One can easily sustain positive rewards in equilibrium among players in the initial periods by introducing incomplete information as in Kreps and Wilson (1982, see also Healy, 2007). The insight of this literature is that it is sufficient for some players to believe in reciprocation as this will cause others to mimic their behavior before an endgame effect sets in. Trust in equilibrium can be predicted in all periods when players care about the outcomes of their opponents such as motivated by concerns for fairness. Similarly one could introduce uncertainty about the number of periods to be played. While the intuition is clear, the modeling details and analysis are space consuming and

[4]As in Burks et al. (2003), the participants in our experiment can take at the same time the roles of sender and receiver. Moreover the dynamic structure of our experiment is similar to the one of Cochard et al. (2004). Differently than both these papers, however, our setting is multilateral, not bilateral.

[5]In our framework, lack of trust can emerge either because senders assign a small probability to the event that their partner will reciprocate or because senders are risk averse. The distinction between these two possible reasons for not trusting others is outside the scope of this paper.

distracting. Hence, we do not include an analytic model in this paper.

3 Results

We organize our analysis as follows. At first we measure overall trust and trustworthiness and consider how these change over time. Next we identify differences in behavior and in success between the participants. We then search for similarities among our participants, similarities that can be quantified in terms of personal information and that translate into overall payoffs attained in the experiment. Similarities are found within regions of nationality, belonging to northern or to southern Europe. We then analyze whether these similarities in terms of success result from systematic differences in terms of behavior. An investigation of how past experience influences future choices rounds off the picture.

We present our analysis with the help of graphs and, to better understand the significance of our results, with the help of regression models with mixed effects. In the regression models we always include a random effect for the matching group and, whenever we look at data with more than one observation per participant, a random effect for the participant. Standard errors, p-values, and confidence intervals are always based on parametric bootstraps with 1000 replications. Given our small subject pool at the EUI we have to work with only 110 participants and only 4 sessions. This rules out more conservative statistical methods. Nevertheless, we do not see any obvious correlation structure in the data that would be neglected by our mixed effects approach.

In our regressions sessions are indexed with s, participants are indexed with i, and j is the period number. To simplify notation we do not write indices $_{ij}$ for variables in the regression equations. Random effects are ϵ_s, ϵ_i and ϵ_{ij} for session s, participant i, and participant i in period j, respectively. We assume that ϵ_s, ϵ_i, and ϵ_{ij} are independent and follow a normal distribution with mean zero. With this specification we allow behavior of the same participant in different repetitions to be correlated as well as different participants from the same session to be correlated.

3.1 Population Behavior

The experiment is embedded in an intricate and flexible environment to allow participants maximum freedom to initiate, build and break relationships with others. In particular we purposely allow more flexibility than in standard trust game experiments. We adapt the steps of our analysis to this special design. To start we ignore personal information and investigate simple measures of behavior. Focus is on the most basic indicators trust and trustworthiness. In this literature, *trust* is measured by the amount of tokens t^S that are

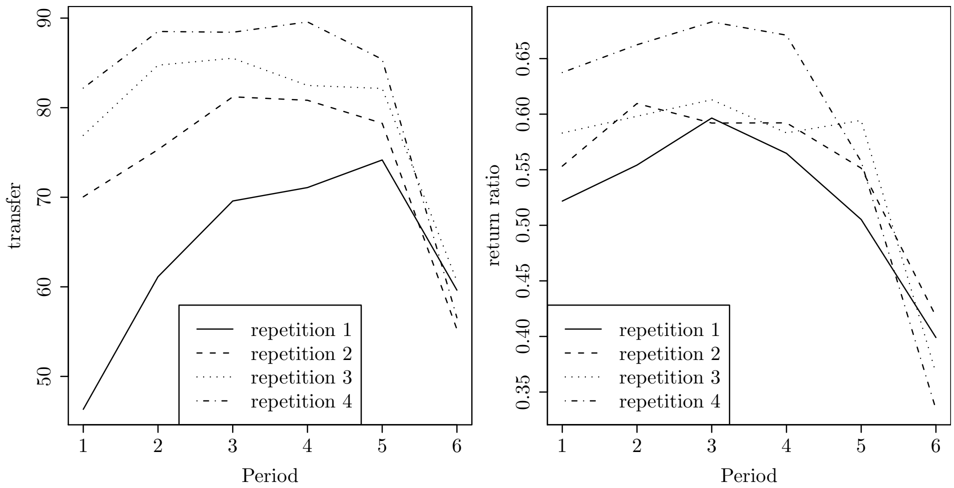


Figure 1: Transfers and returns over time

sent, *trustworthiness* is identified with the *return ratio*. The return ratio, denoted by ρ, is defined only for $t^S > 0$ and is given by the amount of tokens r returned divided by the total amount of tokens received, so $\rho = r/\left(3t^S\right)$.

We compute average trust and average trustworthiness in each period and in each repetition and plot this in Figure 1. Average trust is identified by the average amount of tokens sent. Average trustworthiness is identified by the average return ratio for those interactions where a strictly positive amount of tokens was sent. As the same participant may receive multiple transfers in the same period, the return ratio is first computed separately for each strictly positive transfer before these are averaged.

The figure confirms point 1 we made in the introduction. Overall we observe high degrees of trust and trustworthiness; average trust is equal to 74.4, average trustworthiness is equal to 0.567. These levels of trust and trustworthiness are higher than those observed in one shot trust game experiments (e.g. see Berg et al. 1995, Cochard et al. 2004) and in one shot settings where participants play both roles (Burks et al. 2003), but are similar to those found in previous repeated trust games (Cochard et al. 2004, King-Casas et al. 2005 and first rounds of Engle-Warnick and Slonim, 2004). The fact that participants are in an environment where partners are not preassigned seems to increase the amount of trust as found by Slonim and Garbarino (2008). Within each repetition we find a very similar pattern. Trust and trustworthiness increase at the beginning and decrease at the end of each repetition. In particular the first and the last period are special. The decrease at the end of a repetition is consistent with the findings by Cochard et al. (2004). Similar to Engle-Warnick and Slonim (2004) participants are rematched after every repetition of six periods. It seems that being rematched with new participants at the beginning of each repetition makes participants consider each repetition separately. However, while we observe an increase in trust up to the last period, in the experiment of Engle-Warnick and Slonim (2004), who have 50 repetitions, average trust decreased steadily and substantially within each repetition. Comparing behavior between repetitions we observe an increase in trust and trustworthiness as the experiment progresses.

We assess the significance of our above observations with the help of two mixed effects regressions. Results are shown in the appendix B.1. Marginal trustworthiness in the first five periods is larger than 1/3, i.e. the sender gains in expectation from each additional token she sends. The estimation results confirm that trust increases during the initial stage of a repetition and significantly decreases at the end. There is no significant evidence that trustworthiness differs in the period 1 from periods 2-5 but we find significantly decreasing trustworthiness in the last period.

3.2 Individual Heterogeneity

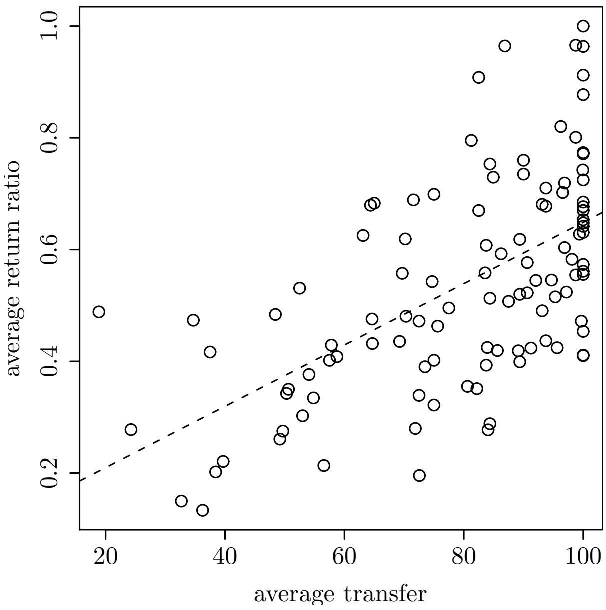
Next we wish to asses whether there are systematic differences between participants. Figure 2 plots average transfers (trust) and average of return ratios (trustworthiness) of each of the participants. As we observed in figure 1 that behavior in the first and last period of each repetition differs from the periods in the middle[6] we consider here only periods 2 to 5.[7]

Note that trust and trustworthiness are correlated across participants. The squared correlation coefficient is 0.347. To test for significance we use a mixed effects model with a random effect for the session and find a p-value below 0.0001. There is mixed evidence in the literature on whether there is such a relationship in one shot games. While some find no such effect (Berg et al. 1995, Willinger et al. 2003, Csukás et al. 2008), others do (Bohnet and Greig 2006). In particular, Burks et al. (2003), Altmann et al. (2007) and Ananish and Gangadharan (2007) find such a positive correlation within participants. Following Figure 2 we observe different degrees of trust and trustworthiness among participants but a systematic relationship between trust and trustworthiness. This allows us to conclude that there are systematic differences between participants (as opposed to differences that stem from randomness).

We now investigate to what extent differences in behavior substantiate in measurable differences in success where success will be measured by payoffs in the experiment. In Figure 3 we plot the average payoffs of the participants and observe a large variation.

[6] This finding is confirmed in the appendix B.1.

[7] We determine the average of individual return ratios as mean($r_{ij}/(3t_{ij})$). Note that this mean disregards observations where $t_{ij} = 0$ as the return ratio is only defined for strictly positive transfers.



Each dot shows averages for one participant.

Figure 2: Average transfer and average return ratio (without periods 1 and 6)

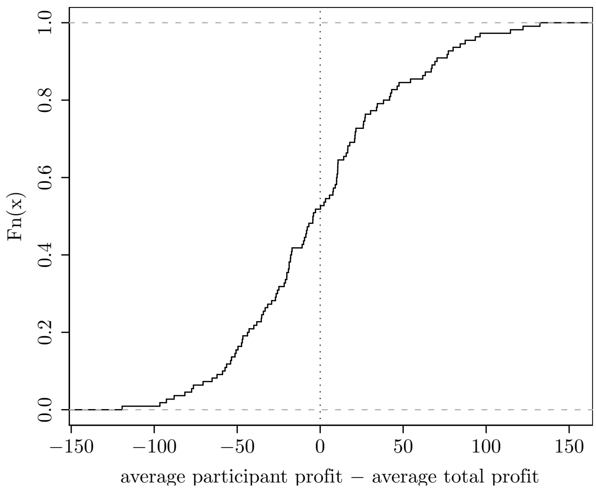


Figure 3: Average participant profit - average total profit

To test whether average payoffs are actually different we estimate

$$\pi_{ij} = \beta_i \cdot d_i + \epsilon_s + \epsilon_i + \epsilon_{ij}$$

where π_{ij} is the average payoff of participant i in the period and repetition indexed by j and d_i is a dummy for each participant i. We can reject the null hypothesis that all participants have the same average payoffs, the p value associated to β_i being independent of i (derived using an F-test) is below 0.0001. Again we can conclude that we find evidence of systematic differences between participants, this time in terms of success.

We expected to find significant differences between participants. To investigate whether or not these differences are systematic, in the sense that they are related to personal characteristics, is the objective of the next section.

3.3 Differences in Success Across Personal Characteristics

We wish to understand whether there is a connection between the above documented differences between participants and their personal characteristics. We will not investigate all possible differences, instead we focus on an important dimension, namely success. Success of participant i will be measured as the average payoff π_i obtained in the experiment.

We now investigate the possibility that observed heterogeneity is related to the personal characteristics that we have collected and made available to participants within their group. These characteristics are age, gender, nationality and number of siblings. Age will be considered as a continuous variable in the regressions. For number of siblings we consider three categories: no, one and more than one sibling. Regarding nationality we compare two regions of origin that we call North and South. For this we take the country latitudes as identified by the CIA database (2003) and include in the category North and South all those countries with latitude above and below the mean. Latitudes and assigned categories are shown in Table 14 in the appendix. The objective of this categorization is to capture commonplace regional differences within Europe. The assignment should be salient and simple and thus as impartial as possible. These concerns lead us to use the CIA data base and the mean latitude of 47.93 as the divider.[8] The robustness of our results to this categorization will also be investigated.

Let us first have a look at Figure 4. The three graphs in this figure show the cumulative distribution of average payoffs (per participant) conditional on region of origin (North/South), gender (male/female) and number of siblings (none, one, and more than one sibling). The only graph that shows visible differences is the one on the left which illustrates the dependency on the region of origin.

To identify significant differences we regress average payoffs π_i on region of origin, gender, age and number of siblings. The following mixed effects model is estimated:

$$\begin{aligned} \pi_i &= \beta_1 + \beta_{d_N} \cdot d_N + \beta_{d_{♂}} \cdot d_{♂} + \beta_A \cdot A \\ &\quad + \beta_{d_{S_1}} \cdot d_{S_1} + \beta_{d_{S_2}} \cdot d_{S_2} + \epsilon_s + \epsilon_i \ . \end{aligned} \quad (1)$$

[8]The median latitude is with 47.2 slightly smaller than the latitude of Austria. The advantage of using the mean is that no country has exactly the mean latitude, hence, we can classify all countries. With the median as the dividing line the classification of Austria would be ambiguous.

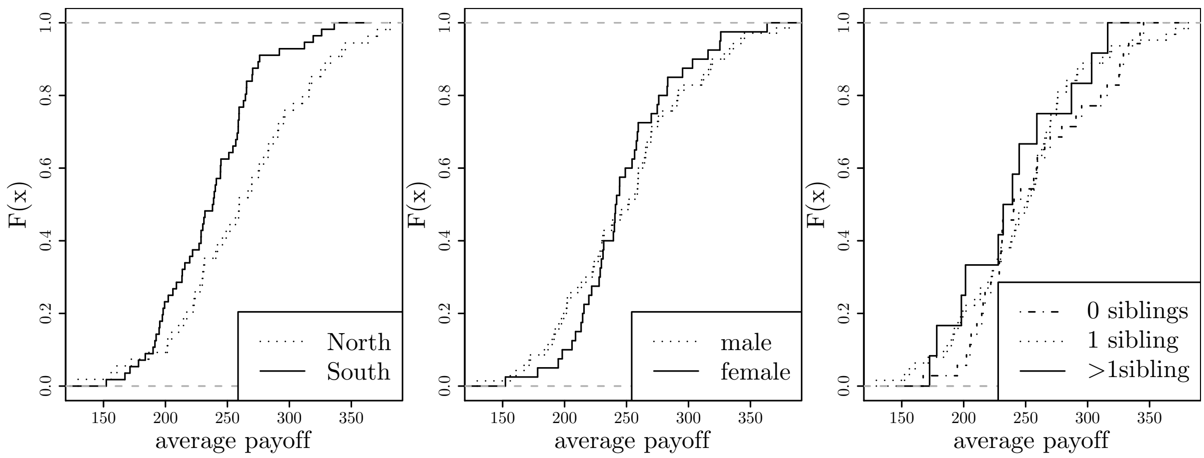


Figure 4: Cumulative distribution of payoffs

	β	σ	t	p value	95% conf	interval	pmvd
1	191	47.4	4.03	0.0001	97.1	285	
d_N	25.5	9.68	2.64	0.0097	6.33	44.7	0.804
$d_{♂}$	-2.85	9.96	-0.286	0.7756	-22.6	16.9	0.008
A	1.25	1.67	0.746	0.4571	-2.07	4.57	0.072
d_{S_1}	12.3	15.7	0.781	0.4368	-18.9	43.5	0.042
d_{S_2}	16.3	16.8	0.973	0.3331	-17	49.6	0.074

Table 1: Determinants of success — estimation of equation (1) payoff π

The dummy d_N is one for participants with a nationality of North (as defined above), the dummy $d_{♂}$ is one for male participants, and A is the age of the participant. d_{S_1} is one for participants with exactly one sibling, d_{S_2} is one for participants with more than one sibling.

Results are shown in Table 1. In addition to reporting p values and confidence intervals we include the proportional marginal value decomposition (short, pmvd) as a measure of how much each variable contributes to the variance of our dependent variable. pmvd was proposed by Feldman (2005), we use the implementation of Grömping (2007). The only significant variable is the region of origin d_N. Controlling for other personal characteristics we find that those from North earn on average 25.54 more tokens per period. This variable is also the main contributor to model performance (pmvd=0.804). This difference is very similar to the one observed without controlling for personal characteristics where we find that those from North earn on average 262 while those from South earn on average 236 tokens per period.

Other experiments do find a significant relation between age and trust and trustworthiness. Bellemare and Kröger (2007) find a hump-shaped relation between age and trust

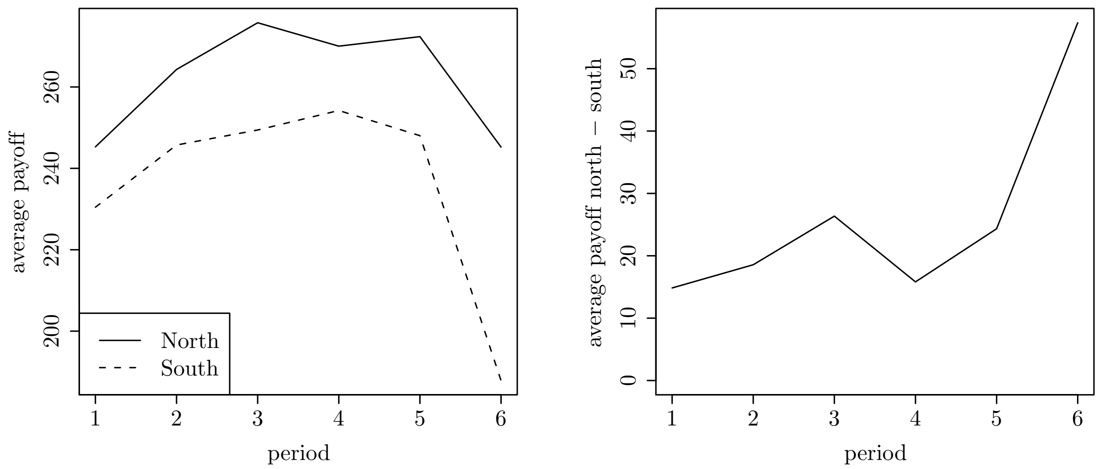


Figure 5: Payoff comparison north and south, per period

and a u-shaped relation between age and trustworthiness. Sutter and Kocher (2007) find a similar relation between age and trust and increasing trustworthiness with increasing age. If trust and trustworthiness depend on age, why isn't age significant in equation (1)? Here we should keep in mind that in our population of EUI students the variability of age is small. The interquartile range of age is only 5 years. This might simply be too small to find a significant effect.

What we can establish in this section is that participants from the North earn significantly more than those from South. In section B.2 of the appendix we present results of two exercises that confirm the robustness of these results.

3.4 A Closer Look at the Success of North

Above we investigated overall success as we only wish to identify differences that matter in this respect. At the same time, success will not be constant over time. In particular, as observed in our analysis of trust and trustworthiness we found substantial differences in behavior between periods. Thus, we expect success to also differ across periods. To highlight these kind of differences, and thus to make a bridge to our more detailed analysis we investigate whether there is a trend across periods and whether regional membership plays a role. Consider first the descriptive statistics in Figure 5.

The left graph in Figure 5 shows the change in success across periods P separately for the two regions North and South. We observe a sharp decline in success in period 6. The graph on the right of Figure 5 shows the difference in success between North and South across periods. We find a constant advantage of North up to period 5 that turns into a substantial advantage in period 6. To support this finding we present in section B.3 of the appendix estimation results from a mixed effects model. While we find no significant change over time in the success of South we observe a significant increase in success of North across periods. The average drop in success in period 6 is strongly significant and substantial.

3.5 Differences in Behavior Across Regions

Above we found systematic differences in success only in terms of the region of origin. We now search for differences in levels of trust or trustworthiness between these two regions that can help understand their differences in success. Given the statistical irrelevance of the other personal characteristics that emerged from the analysis of the previous section, these characteristics are now dropped to save degrees of freedom.

Recall the behavioral regularities we observed for the entire pool of participants. Trust is at a high level and increases between repetitions. Within each repetition the willingness to trust shows a dynamic pattern as there is both caution at the beginning and at the end. Trustworthiness is generally high except for the last period of each repetition where it is substantially lower. Trustworthiness increases slightly between repetitions.

In the following regression we explain the amount of tokens sent, t^S_{ij}, and the amount of tokens returned, r^S_{ij}. Here, i refers to the identity of the participant and j is a *period number* that uniquely identifies the period and the repetition, so $j = 1, .., 24$. The sum of transfers received by participant i in the period numbered j is denoted by t^R_{ij}. So t^R_{ij} is the tripled amount of tokens sent by other players to participant i in the period numbered j. The coefficient of t^R_{ij} in the regression can be considered as a measure of marginal trustworthiness. Since behavior might change over time we include the repetition $T \in \{1, 2, 3, 4\}$ of the experiment

Let d_N be the dummy that is equal to 1 if the participant belongs to North. Index N indicates interactions with North, so for instance $d_{1N} = d_1 \cdot d_N$, $T_N = T \cdot d_N$ and

	β	σ	t	p value	95% conf	interval
d_1	47.3	3.91	12.1	0.0000	39.6	54.9
d_{2-5}	60.7	3.55	17.1	0.0000	53.7	67.6
d_6	43.5	3.82	11.4	0.0000	36	51
T	6.1	0.609	10	0.0000	4.9	7.29
d_{1N}	13.6	4.51	3.02	0.0025	4.78	22.5
d_{2-5N}	8.84	3.87	2.29	0.0224	1.25	16.4
d_{6N}	-0.715	4.42	-0.162	0.8715	-9.38	7.95
T_N	-0.274	0.867	-0.316	0.7520	-1.97	1.43

Table 2: Determinants of trust — estimation of equation (3), t^{S}

$t^R_{6N} = t^R_6 \cdot d_N = t^R \cdot d_6 \cdot d_N$. Given

$$\begin{aligned} X &\equiv \beta_{d_1} \cdot d_1 + \beta_{d_{2-5}} \cdot d_{2-5} + \beta_{d_6} \cdot d_6 + \beta_T \cdot T \\ &\quad + \beta_{d_{1N}} \cdot d_{1N} + \beta_{d_{2-5N}} \cdot d_{2-5N} + \beta_{d_{6N}} \cdot d_{6N} + \beta_{T_N} \cdot T_N \qquad (2) \end{aligned}$$

then we estimate the following mixed effects models

$$t^{\mathrm{S}} = X + \epsilon_s + \epsilon_i + \epsilon_{ij} \qquad (3)$$

$$\begin{aligned} r^{\mathrm{S}} &= X + \beta_{t^R_1} \cdot t^R_1 + \beta_{t^R_{2-5}} \cdot t^R_{2-5} + \beta_{t^R_6} \cdot t^R_6 + \beta_{t^R_T} \cdot t^R_T + \beta_{t^R_{1N}} \cdot t^R_{1N} + \\ &\quad + \beta_{t^R_{2-5N}} \cdot t^R_{2-5N} + \beta_{t^R_{6N}} \cdot t^R_{6N} + \beta_{t^R_{TN}} \cdot t^R_{TN} + \epsilon_s + \epsilon_i + \epsilon_{ij} \qquad (4) \end{aligned}$$

Equations (3) and (4) are actually an extension of regressions (7) and (8) which we mentioned in our discussion of figure 1 in section 3.1. The detailed results for these regressions can be found in section B.1 of the appendix.

Information about regional membership gives insights into differences in behavior. Consider trust (see Table 2). Comparing to trust of South we find that trust of North is significantly higher in periods 1-5 while it is not significantly different in period 6 nor is there a significant difference in the change between repetitions.

Consider now tokens returned (see Table 3). We find a close linear relationship between tokens returned and tokens received as the coefficients β_{d_1}, $\beta_{d_{2-5}}$, β_{d_6}, $\beta_{d_{1N}}$, $\beta_{d_{2-5N}}$ and $\beta_{d_{6N}}$ are insignificant. Only β_T is significant but β_{T_N} is insignificant. We find that tokens returned mainly depend on the amount of tokens received. So we now look at the coefficients $\beta_{t^R_\cdot}$ which are associated to the amount of tokens received and which measure marginal trustworthiness. Comparing to marginal trustworthiness of South, we find marginal trustworthiness of North to be similar to South in period 1 ($\beta_{t^R_{1N}} = -0.0517$) but lower in periods 2-6 ($\beta_{t^R_{2-5N}} = -0.105$, $\beta_{t^R_{6N}} = -0.166$). Between repetitions we find that trustworthiness of North changes significantly more than that of South ($\beta_{t^R_{TN}}$ is significantly positive). The estimated increase in marginal trustworthiness among North,

	β	σ	t	p value	95% conf	interval
d_1	-5.97	15.1	-0.395	0.6926	-35.6	23.6
d_{2-5}	-8.98	13.7	-0.656	0.5118	-35.8	17.8
d_6	-12.1	14.8	-0.815	0.4150	-41.2	17
T	5.74	2.55	2.25	0.0244	0.741	10.7
d_{1N}	12.2	14.9	0.817	0.4138	-17.1	41.5
d_{2-5N}	16.2	12.9	1.25	0.2103	-9.14	41.5
d_{6N}	15.5	15.5	0.998	0.3183	-14.9	45.9
T_N	-6.71	3.74	-1.79	0.0730	-14	0.625
t_1^{R}	0.543	0.0338	16.1	0.0000	0.476	0.609
t_{2-5}^{R}	0.596	0.0258	23.2	0.0000	0.546	0.647
t_6^{R}	0.422	0.0353	12	0.0000	0.353	0.491
t_T^{R}	-0.0109	0.00909	-1.2	0.2303	-0.0287	0.00691
t_{1N}^{R}	-0.0517	0.0458	-1.13	0.2589	-0.142	0.0381
t_{2-5N}^{R}	-0.105	0.0347	-3.03	0.0024	-0.173	-0.0372
t_{6N}^{R}	-0.166	0.0465	-3.58	0.0004	-0.258	-0.0752
t_{TN}^{R}	0.0467	0.012	3.89	0.0001	0.0232	0.0702

Table 3: Determinants of trustworthiness — estimation of equation (4), r^{S}

$\beta_{t^R_{TN}} + \beta_{t^R_T} = 0.0358, \mathrm{CI}_{95} = [0.0193, 0.0522]$ is small, though significantly positive (CI denotes the 95% confidence interval based on a parametric boostrap. Details of all the calculations in this section can be found in section B.4 of the appendix).

We note that the marginal trustworthiness of North is lower in periods 2-6 but at the same time the intercept is positive albeit insignificant. To understand the connection between these two effects we estimate the return ratio of North. Consider period 1. Following Table 10 we estimate the average transfer in period 1 to be 68.9. Following Table 3 North returns on average $124, \mathrm{CI}_{95} = [96.4, 151]$ in period 1. Hence, the return ratio is $0.599, \mathrm{CI}_{95} = [0.48, 0.717]$. Similarly we estimate the return ratio for South in period 1 to be $0.556, \mathrm{CI}_{95} = [0.432, 0.68]$. For periods 2-5 we find analogously an estimated return ratio of North and of South equal to $0.601, \mathrm{CI}_{95} = [0.502, 0.699]$ and $0.592, \mathrm{CI}_{95} = [0.493, 0.69]$, respectively. For period 6 we find $0.351, \mathrm{CI}_{95} = [0.203, 0.499]$ and $0.408, \mathrm{CI}_{95} = [0.262, 0.555]$, respectively. So trustworthiness of North in terms of return ratio is not lower than that of South in periods 1-5, only in period 6 we estimate that North has a slightly (though not significantly) lower trustworthiness than South.

To sum up, we find the differences between North and South are consistently in the direction of generating more tokens for North. Consider for instance differences in trust. In Section 3.1 we found for periods 1-5 that trust is rewarded since trustworthiness is significantly above 1/3 while this is not so in period 6. So North trusts more when it pays off (periods 1-5). In terms of differences in trustworthiness, North differs by

	β	σ	t	p value	95% conf	interval
d_1	164	25.1	6.52	0.0000	115	213
d_{2-5}	187	21.6	8.69	0.0000	145	229
d_6	115	24.6	4.69	0.0000	67	163
T	13.7	4.92	2.79	0.0053	4.09	23.4
d_{1N}	-3.86	31.9	-0.121	0.9038	-66.4	58.7
d_{2-5N}	15.8	25.9	0.61	0.5422	-35	66.6
d_{6N}	28.8	31.1	0.926	0.3545	-32.2	89.7
T_N	8.44	7	1.21	0.2280	-5.29	22.2

Table 4: Trust of others — estimation of equation (5), t^{R}

returning significantly less than South at the margin in periods 2-6. In terms of return ratio there are no big differences. The only case where North is generating less tokens than South, and possibly is acting strategically, is across repetitions where North is increasing trustworthiness.

3.6 Observability of Regions

We now investigate to what degree these regional differences are perceived by others. Or in other words, to what extent does own nationality influence behavior of others? This is important as the differences in success could be driven by discrimination based on nationality. We consider both trust and trustworthiness of others. We replace t^S by t^R in (3) to obtain (5). We also replace r^S by r^R and t^R by t^S in (4) to obtain (6). r^{R} is the returned amount a player receives. The only other change is that we now have to condition on observed age as the true age is not observed by others. So we estimate

$$t^{\mathrm{R}} = X + \epsilon_s + \epsilon_i + \epsilon_{ij} \tag{5}$$

$$\begin{aligned} r^{\mathrm{R}} = {} & X + \beta_{t_1^S} \cdot t_1^S + \beta_{t_{2-5}^S} \cdot t_{2-5}^S + \beta_{t_6^S} \cdot t_6^S + \beta_{t_T^S} \cdot t_T^S + \beta_{t_{1N}^S} \cdot t_{1N}^S + \\ & + \beta_{t_{2-5N}^S} \cdot t_{2-5N}^S + \beta_{t_{6N}^S} \cdot t_{6N}^S + \beta_{t_{TN}^S} \cdot t_{TN}^S + \epsilon_s + \epsilon_i + \epsilon_{ij} \end{aligned} \tag{6}$$

Consider for instance the dummy d_N. In (5) it describes how transfer of others to a participant from North depends on the fact that this participant belongs to North. In (6) it describes how the tokens returned depend on the fact that the sender belongs to region North. Looking at Tables 4 and 5 we see that there is no significant evidence that behavior depends on the information provided about regional membership.

The remainder of the section attempts to quantify if the observed differences in transfers and returns between regions are enough to explain the differences in payoffs. Following Table 1 participants from the North earn 25.5 more tokens than South. How do they do

	β	σ	t	p value	95% conf	interval
d_1	-8.33	19.6	-0.425	0.6706	-46.7	30.1
d_{2-5}	-16.5	18	-0.915	0.3604	-51.7	18.8
d_6	-8.94	19.2	-0.465	0.6418	-46.6	28.7
T	0.469	4.3	0.109	0.9131	-7.95	8.89
d_{1N}	14.8	20.4	0.727	0.4671	-25.1	54.8
d_{2-5N}	13.5	17.2	0.787	0.4313	-20.2	47.2
d_{6N}	12	20.4	0.588	0.5567	-28	52
T_N	-0.571	6.34	-0.0901	0.9282	-13	11.9
t_1	1.62	0.198	8.18	0.0000	1.23	2.01
t_{2-6}	1.82	0.154	11.8	0.0000	1.52	2.12
t_6	0.972	0.181	5.38	0.0000	0.618	1.33
t_T	0.0566	0.0551	1.03	0.3040	-0.0514	0.165
t_{1N}	-0.0871	0.275	-0.317	0.7516	-0.627	0.452
t_{2-5N}	-0.147	0.216	-0.681	0.4961	-0.57	0.276
t_{6N}	0.0949	0.253	0.376	0.7071	-0.4	0.59
t_{TN}	0.00627	0.0757	0.0828	0.9340	-0.142	0.155

Table 5: Success of trust — estimation of equation (6), r^{R}

this? Above we observed differences between North and South that translated into generating more payoffs. Together with Table 2 we can determine that the difference in transfer between North and South in periods 2-5 generates a net gain of $5.08, \mathrm{CI}_{95} = [0.707, 9.46]$ more tokens for North (Details of the calculations in this section can be found in section B.5 of the appendix). So while North transfers more than South, this does not translate into substantially higher payoffs. A second channel to explain differences in payoffs between North and South is the number of tokens returned. Consider again periods 2-5. Together with Table 3 we estimate that in periods 2-5 North earn due to the difference in their return behavior $4.35, \mathrm{CI}_{95} = [-16.4, 25.1]$ more tokens on average than South. The combined effect for periods 2-5 is, thus, $5.08 + 4.35 = 9.43, \mathrm{CI}_{95} = [-11.8, 30.6]$. Similarly, we calculate the effect for period 1 as $3.21, \mathrm{CI}_{95} = [-21.1, 27.5]$ and for period 6 as $6.05, \mathrm{CI}_{95} = [-19.6, 31.7]$. The average effect for all 6 periods is, hence, $7.83, \mathrm{CI}_{95} = [-13, 28.7]$ which is considerably, although not significantly, below the estimated difference in payoffs of 25.5 in Table 1. We conclude that differences found in trust and trustworthiness generate more payoffs for North but are hardly enough to explain total payoff differences. The reason is that so far we have ignored how differences in trust and trustworthiness possibly influence whether or not one receives transfers in future periods. This is the objective of the next section.

3.7 Selection of Partners

We now wish to come back to point 2 of our introduction and to investigate how partners are selected and how these selections depends on previous outcomes. We are specifically interested in (i) how differences in trust and trustworthiness influence the probability of being selected in the next period, (ii) uncovering regional similarities and differences in the way participants condition behavior on past outcomes, and (iii) whether there is discrimination between regions in the selection of partners. We focus on the influence of experiences from the last period on the next selection and consider separately past outcomes resulting from making an own selection and past outcomes resulting from being selected. The behavioral model that we have in mind is partially based on reinforcement. One can imagine that a participant could unconditionally be more likely to select the same participant again, for instance because of inertia, or alternatively, in order to build trust. One can imagine that the selection of the same participant again could be conditional on the amount of tokens returned, hence on the revealed trustworthiness of the participant chosen. This would be in line with reinforcement. Note that there is another important ingredient to our design, namely that one can also be selected by another participant and, thus, receive tokens. When begin selected, one could unconditionally choose to reward this behavior by selecting this participant in the next period. In line with reinforcement one could condition such a reward on the amount of tokens received. Changes in tokens sent or tokens returned as a function of past experience is not considered here, the main dynamic phenomena that depended on the period were documented above.

3.7.1 Selection of Partners From North and South

To understand the role of past information on the dynamic choice of partners we develop a conditional logit model. Results are shown in Table 6, these estimates are based on the conditional logit model described in the appendix C.

d_{ikp-1} is a dummy variable that describes whether or not participant i transferred tokens to participant k in the previous period. Consider the first column of this table. The coefficient of this dummy is equal to 3.43 and is strongly significantly different from 1. This is statistical evidence that participants tend to select whom they selected in the previous period, the odds to select again the same as opposed to selecting someone different is estimated to be 3.43. In the specification underlying the first column we also included d_{kip-1} which is an indicator of whether participant i was selected by participant k in the previous period. Again, the coefficient is strongly significantly different from 1 with a similar magnitude as d_{ikp-1}. Participants tend to trust those who trusted them.

	$\hat{\beta}$	s.e.	$\hat{\beta}$	s.e.	$\hat{\beta}$	s.e.
d_{ikp-1}	3.43	0.17***	3.42	0.17***	1.75	0.12***
d_{kip-1}	4.47	0.30***	4.5	0.30***	2.93	0.27***
d_{NN}			1.05	0.08	0.93	0.08
d_{SS}			0.81	0.07***	0.9	0.08
$d_{ikp-1} \cdot 1\{r_{ikp-1} \geq \text{med}(r)\}$					3.77	0.39***
$d_{kip-1} \cdot 1\{t_{kip-1} \geq \text{med}(t)\}$					2.37	0.27***
fraction of correct predictions	0.62		0.61		0.63	

Note: d_{ikp-1} denotes the event that i has chosen k in the previous period, and d_{kip-1} means that k has chosen i in the previous period. "$\wedge$" is the logical "and" operator, and $1\{...\}$ is the indicator function which takes value one if the expression inside the parenthesis is true. "med()" is the median of the variable. Reported values are odds ratios, standard errors in parenthesis. *, ** , *** denote significance equal to the 10, 5 and 1 percent level of a test that the odds ratio is one. Control variables included are: age and siblings of all 4 players, gender and nationality of i interacted with the attributes of k. Correct predictions indicates the share of observations in which the highest estimated probability $\hat{p}_k$ coincides with the actual choice. See the appendix C for further details.

Table 6: Dynamic choice model, all periods 2-6

In the second column of Table 6 we add variables to investigate whether discrimination between regions plays a role in the selection process. d_{NN} is a dummy that indicates for participants from the North whether they are selecting in the current period a participant from the same region, namely North. Similarly, d_{SS} is equal to 1 if a participant from South is selecting a participant from South. We find that d_{SS} is significantly different from 1, the estimate is 0.81 which is less than 1. Thus, we find strongly significant evidence that those from South discriminate against participants from the same region. We do not find any evidence of discrimination on the part of participants from North.

In the third column of Table 6 we wish to investigate to what extent selection not only depends on previous selections but also on tokens transferred, thus anwering point 3 from the introduction. For this we interact the dummies of the first column with an indicator for whether the transfers to or the return ratios received from k in the previous period were larger than the corresponding median levels in the sample. The estimated odds ratios for these interactions are larger than 1 and highly significant as well. The first one (3.77) indicates that if i selected k in the previous period then i is more likely to select k again if k previously returned more than the median in the sample. Thus, as expected, higher trustworthiness is rewarded with a higher likelihood to be selected. The second interaction is, however, more interesting for our purposes. It indicates that if k selected i in period 1, then i is more likely to select k in period 2 if the previous transfer from k was larger than the median in the sample. Thus, also a higher degree of trust from others is rewarded with trust. Note that the coefficients connected to discrimination that we add in column two, and which were significant there, now are no longer significantly different from 1. In fact,

	$\hat{\beta}$	s.e.	$\hat{\beta}$	s.e.	$\hat{\beta}$	s.e.
d_{ikp-1}	3.34	0.37***	3.34	0.37***	1.97	0.3***
d_{kip-1}	6.67	1.02***	6.7	1.03***	4.44	0.82***
d_{NN}			1.43	0.25***	1.34	0.24***
d_{SS}			0.84	0.16	0.99	0.2
$d_{ikp-1} \cdot 1\{r_{ikp-1} \geq \text{med}(r)\}$					3.46	0.81***
$d_{kip-1} \cdot 1\{t_{kip-1} \geq \text{med}(t)\}$					3.17	0.83***
fraction of correct predictions	0.67		0.65		0.65	

Table 7: Dynamic choice model, period 2

even their estimates are close to 1. Thus, we were picking up behavior that was specific for those belonging to North and not uncovering discrimination per se. For instance, the change in the significance of d_{SS} could be evidence that it was not that South was discriminating in favor of North because of the region. Instead South was discriminating because those from North tend to transfer more. Of course this interpretation has to be taken with a grain of salt as North also tend to be less trustworthy.

To summarize, we find no evidence of discrimination in terms of behavior that can only be explained by the observed nationality, but instead an intricate dependency of behavior on past outcomes. There is evidence of unconditional behavior, to select the same again or to select reciprocally. There is evidence of conditional behavior, to select the trustworthy or to select the trusting.

3.7.2 Changes in the Selection Process

We combine the above insight with our observed differences between North and South. As North transfers more tokens than South in periods 1-5, thus, North is more likely to be selected in the next period. Notice that payoffs increase with the likelihood of being selected. So higher trust of North is amplified via reinforcement. In terms of return ratio we found no substantial differences in periods 1-5. Hence, trustworthiness does not add to explaining the differences in success in this dynamic story.

Next we consider whether our above insights remain when we consider the beginning and the end of each repetition separately. This is motivated from our previous findings that behavior in periods 1 and 6 is different from that periods 2-5. Hence, we now rerun the same three specifications shown in Table 7 separately for period 2 and period 6. Evidence is given in Tables 7 and 8.

In period 2 we now find that North have a significant preference for North, the magnitude is roughly that North is 40% more likely to choose North than to choose South. This evidence of discrimination of North in favor of North persists when we add the in-

	$\hat{\beta}$	s.e.	$\hat{\beta}$	s.e.	$\hat{\beta}$	s.e.
d_{ikp-1}	3.65	0.44^{***}	3.67	0.44^{***}	1.81	0.31^{***}
d_{kip-1}	3.04	0.5^{***}	3.04	0.49^{***}	2.04	0.47^{***}
d_{NN}			1.15	0.24	0.97	0.21
d_{SS}			1.09	0.24	1.2	0.27
$d_{ikp-1} \cdot 1\{r_{ikp-1} \geq \text{med}(r)\}$					4.43	1.17^{***}
$d_{kip-1} \cdot 1\{t_{kip-1} \geq \text{med}(t)\}$					2.19	0.61^{***}
fraction of correct predictions	0.61		0.6		0.63	

Table 8: Dynamic choice model, period 6

teractions with the transferred or returned amounts. Of course we may just be capturing behavioral specific effects common to those in North that we have not included. However, this does not seem plausible given our previous finding that the difference in behavior between North and South is similar in period 1 and in periods 2-5. In fact, we found that the difference in behavior changes more dramatically between periods 2-5 and period 6. Yet in Table 8 we now no longer find any evidence of discrimination.

To summarize, we note no substantial difference in how participants condition on past experience between the second and the last period as compared to all periods. Our insights remain. However, we do find that North start out by discriminating more in favor of unconditionally selecting participants from their region and that this effect is no longer visible in period 6. If there is discrimination then our findings show that this only holds for period 2.

We are now ready to compare the dynamic behavior of North and South. In particular we are interested in whether we find more evidence in favor of why North earn more than South. To do this we add the interactions with North to the equation whose estimates are shown in the third column of Table 6. The estimates are shown in Table 9.

Here d_N is a dummy that records whether or not participant i belongs to region N. There remains to be no evidence of discrimination as both d_{NN} and d_{SS} are insignificant. The reinforcement behavior of North differs from that of South in terms of the increased reward of receiving more trust than the median amount. The estimate is equal to 1.71 and is strongly significant. It seems like North are more sensitive to receiving high trust from others. Since high trust is also correlated with high trustworthiness, a selection of others from whom more trust (larger amounts) were received is a good strategy to increase payoffs. Hence, we uncover a further reason based on dynamic behavior why North are more successful than South.

	$\hat{\beta}$	s.e.
d_{ikp-1}	1.67	0.16***
$d_{ikp-1} \cdot d_N$	1.11	0.16
d_{kip-1}	3.28	0.4***
$d_{kip-1} \cdot d_N$	0.78	0.14
$d_{ikp-1} \cdot 1\{r_{ikp-1} \geq \text{med}(r)\}$	4.08	0.6***
$d_{ikp-1} \cdot 1\{r_{ikp-1} \geq \text{med}(r)\} \cdot d_N$	0.85	0.18
$d_{kip-1} \cdot 1\{t_{kip-1} \geq \text{med}(t)\}$	1.83	0.29***
$d_{kip-1} \cdot 1\{t_{kip-1} \geq \text{med}(t)\} \cdot d_N$	1.71	0.39**
d_{NN}	0.92	0.08
d_{SS}	0.88	0.08
fraction of correct predictions	0.63	

Table 9: Dynamic choice model, all periods with interaction

4 Conclusion

Trust and trustworthiness play a crucial role in the selection of partners and in the development of social and economic interactions. But while they are typically thought of as static and time invariant cultural traits of agents, this paper shows that they build up and evolve dynamically with experience in a setting of repeated interactions.

In our dynamic version of the trust game in which participants from different European nationalities can repeatedly select their partners and choose the size of their interactions with them, Northern Europeans emerge with higher payoffs. This is not so much the result of aprioristic taste based discrimination. North are better investors from the outset of the game. Their experience in the game reinforces this trait and give them higher payoffs because they are paid back more and they receive more contacts in the future. Moreover, thanks to the initial larger set of contacts North can later select partners from a larger set of participants, of which they had an opportunity to experience trustworthiness. In other words, trust breeds trust and allows to identify where to find trustworthiness.

These findings suggest that a convergence to two substantially different population equilibria in two societies can emerge from a grain of initial "cultural" difference that reinforces itself in a dynamic trajectory leading to a major divergence of economic success. For example, a small group of individuals endowed with low trust and trustworthiness can cause a snowball effect by which more and more people adopt their standards as trust and trustworthiness pays off less and less. And vice-versa, a small group of high trust individuals can dynamically generate very opposite outcomes.

We have nothing to say on what may determine the initial "grain of cultural difference" between north and south that emerges from our experiment as well as from the evidence of Knack and Keefer's (1997) and Guiso, Sapienza and Zingales (2004b). While this is outside the scope of this paper we suggest two directions here: The first possibility is that these differences emerge merely from an income effect: Assuming that "generosity" and "reciprocity" are luxury (normal) goods, people will tend to "consume" more of them the greater is their income.

Thus, the higher level of income and stage of development in the North during recent history would be responsible for cultural differences regarding trust and trustworthiness, reflected in our results. The other possible explanation is that initial differences in terms of trust and trustworthiness between South and North of Europe have to do with the different role of the family in these two regions. A similar hypothesis was suggested by Banfield (1958) in his study of life in the southern Italian rural village of Chiaromonte. Banfield describes the behavior of this population as governed by the code of "amoral familism", according to which moral principles are regarded as irrelevant by residents of the village when they deal with non-family members. Also at the more general European level, there is evidence suggesting that in both social and economic activities the family plays a much greater role in the South than in the North. With family ties less intensive in the North, people in this region rely on networking outside the family more than people in the South.[9] Trust and trustworthiness outside the family is thus more crucial for social and economic success in the North. Related to this interpretation is also the work of Putnam (1993) who emphasizes the role of social capital in explaining the economic backwardness of Southern Italy.

Whatever the reason of the "initial grain of cultural difference", the main contribution of our paper is to show that it can be reinforced by experience, multiplying its effects and leading to differences in success which are much larger than what one would consider possible in a static environment. This has important implications for the hesitant process of European unification. It is possible that repeated reciprocal experiences among Europeans leads to divergence of outcomes and not to integration, in which case the Union would have to carefully study strategies to invert the divergence of these trends.

References

Acemoglou, D., Johnson, S., Robinson, J.A. (2001). The colonial origins of comparative development: An empirical investigation. American Economic Review, 91, 1369-1401.

[9]See, for example, Bentolila and Ichino (2007) and their references. A similar relationship between family ties and attitudes towards trust and trustworthiness is documented for China by Weber and Hsee (1998).

Altmann, S., Dohmen, T., Wibral, M. (2007). Do the reciprocal trust less? IZA Discussion Paper No. 3010.

Ananish, C., Gangadharan, L. (2007). An experimental analysis of trust and trustworthiness. Southern Economic Journal, 73(4), 959-985.

Arrow, K. (1972). Gifts and exchanges. Philosophy and Public Affairs, 1, 343-362.

Banfield, E.C. (1958). The moral basis of a backward society. New York: The Free Press.

Bellemare, C., Kröger, S. (2007). On representative social capital. European Economic Review, 51, 183–202.

Bentolila, S., Ichino, A. (2007). Unemployment and consumption near and far away from the Mediterranean?. Journal of Population Economics, 21, 255-280.

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

Buchan, N.R., Johnson, E.J., Croson, R.T.A. (2006). Let's get personal: An international examination of the influence of communication, culture and social distance on other regarding preferences. Journal of Economic Behavior & Organization, 60(3), 373-398.

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

Central Intelligence Agency (CIA) (2003). The World Fact Book, http://www.cia.gov/cia/publications/factbook/.

Charness, G., Du, N., Yang, C.-L. (2009). Trust and trustworthiness reputations in an investment game. Mimeo.

Cochard, F., Nguyen Van, P., Willinger, M. (2004). Trusting behavior in a repeated investment game. Journal of Economic Behavior and Organization, 55(1), 31-44.

Csukás, P.F., Kovács, T., Willinger, M. (2008). The determinants of trusting and reciprocal evidence from an intercultural experiment. Journal of Economic Development 33(1), 71-95.

Engle-Warnick, J., Slonim, R.L. (2006). Learning to trust in indefinitely repeated games. Games and Economic Behavior, 54(1), 95-114.

Feldman, B. (2005). Relative importance and value. Kellstadt Graduate School of Business, mimeo, http://www.prismanalytics.com/docs/RelativeImportance.pdf.

Fershtman, C. and Gneezy, U., (2001). Discrimination in a segmented society: An experimental approach. Quaterly Journal of Economics, 116, 351-377.

Fershtman, C., Gneezy, U. Verboven, F. (2005). Discrimination and nepotism: The efficiency of the anonymity rule. Journal of Legal Studies, 34(2), 371-394.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for readymade economic experiments. Experimental Economics, 10(2), 171-178.

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

Glaeser, E.L., La Porta, R. Lopez-de-Silanes, F. Schleifer, A. (2004). Do Institutions cause growth. NBER Working Paper No. W10568.

Glaeser, E.L., Laibson, D.I., Scheinkman, J.A., Soutter, C.L. (2000). Measuring trust. Quarterly Journal of Economics, 115(3), 811-846.

Greif, A. (1994). Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies?. Journal of Political Economy, 102(5), 912-950.

Greig, F., Bohnet, I. (2006). Is there reciprocity in a reciprocal-exchange economy? Evidence of gendered norms from a slum in Nairobi, Kenya. Mimeo, http://ksghome.harvard.edu/~ibohnet/researchpdf/Greig%20Bohnet%2006_01_06.pdf.

Grömping, U. (2007). Estimators of relative importance in linear regression based on variance decomposition. The American Statistician 61, 139-147.

Guiso, L., Sapienza, P., Zingales, L. (2004a). The role of social capital in financial development. American Economic Review, 94, 526-556.

Guiso, L., Sapienza, P., Zingales, L. (2004b). Cultural biases in economic exchange, NBER Working Paper No. 1105.

Hall, R.E., Jones, C. (1999). Why do some countries produce much more output per worker than others?. Quarterly Journal of Economics, 114(1), 83-116.

Healy, P.J. (2007). Group reputations, stereotypes, and cooperation in a repeated labor market. American Economic Review, 97, 1751-1773.

Holm, H.J., Danielson, A. (2005). Tropic trust versus nordic trust: Experimental evidence from Tanzania and Sweden. Economic Journal, 115, 505-532.

Holm, H.J., Nystedt, P. (2005). Intra-generational trust—a semi-experimental study of trust among different generations. Journal of Economic Behavior and Organization, 58(3), 403-419.

Kahneman, D., Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. Econometrica, 47, 313-327.

King-Casas, B., Tomlin, D., Anen, C., Camerer, C.F., Quartz, S.R., Read Montague, P., (2005). Getting to know you: Reputation and trust in a two-person economic exchange. Science, 308(5718), 78-83.

Knack, S., Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. Quarterly Journal of Economics, 52, 1251-1287.

Kreps, D.M., Wilson, R. (1982). Reputation and imperfect information. Journal of Economic Theory, 27, 253-279.

Kruskal, W. (1987). Relative importance by averaging over orderings. The American Statistician, 41, 6-10.

La Porta R., Lopez-De-Silanes, F., Schleifer, A., Vishny, R. (1997). Trust in large organizations. American Economic Review Papers and Proceedings, 87, 333-338.

La Porta R., Lopez-De-Silanes, F., Schleifer, A., Vishny, R. (1999). The quality of government. The Journal of Law, Economics and Organization, 15, 222-279.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), Frontiers in Econometrics. New York: Academic Press.

Putnam, R.D. (1993). Making democracy work. Civic traditions in modern Italy. Princeton-NJ: Princeton University Press.

Schotter, A. (1981). The economic theory of social institutions. Cambridge, MA: Cambridge University Press.

Schotter, A., Sopher, B. (2006). Trust and trustworthiness in games: An experimental study of intergenerational advice. Experimental Economics 9, 123-145.

Slonim, R. (2006). Gender selection discrimination: Evidence from a trust game. Case Western Reserve University, mimeo.

Slonim, R., Garbarino, E. (2008). Increases in trust and altruism from partner selection: Experimental evidence. Experimental Econonomics, 11, 134-153.

Sutter, M., Kocher, M.G. (2007). Trust and trustworthiness across different age groups. Games and Economic Behavior, 59, 364–382

Tabellini, G. (2005). Culture and institutions: Economic development in the regions of Europe. Bocconi University, mimeo.

Weber, E.U., Hsee, C.K. (1998). Cross-cultural difference in risk perception, but cross cultural similiarities in attitudes towards perceived risk. Management Science, 44(9), 1205-1217.

Willinger, M., Keser, C., Lohmann, C., Usunier, J.-C. (2003). A comparison of trust and reciprocity between France and Germany: Experimental investigation based on the investment game. Journal of Economic Psychology, 24, 447-466.

Appendix

A Experimental Design

The experiment was conducted using a computerized setup[10] in 4 sessions at the European University Institute near Florence, Italy. Participants were 110 Masters and PhD students from the faculties of Law (30%), History (15%), Social and Political Sciences (23%), and Economics (33%). Participants originated from 15 different European countries. They were between 23 and 36 years old (average: 27.7) and 64% were male. Because it was the first time that experiments were conducted at this place, the pool of participants was not experienced in playing games. For each session a multiple of five participants was recruited (session 0: 10, session 1: 30, session 2: 40, session 3: 30). The profit earned by participants ranged from €24 to €47.90, with an average of €36.34 (s.d. 4.89), including a €5 show-up fee paid to each candidate. Each session lasted for about 2 hours. Participants were recruited via email and were invited to sign up on a web site. Each session took place in 2 to 3 computer labs with 10 to 25 computers each, located in different buildings of the university campus. Upon arrival to an assigned computer lab, participants randomly drew a seat number and an account number. This account number was later used to identify participants for payment, which was organized anonymously. Further to that, the computer labs were prepared using separators to individualize the environment. In each room, a professor of the university monitored the experiment in a discrete way.

Note that at no point in time were participants deceived. Participants could choose how often (max 3 times) they wanted to read through the instructions on the screen.

[10] The z-Tree software is described in Fischbacher (2007)

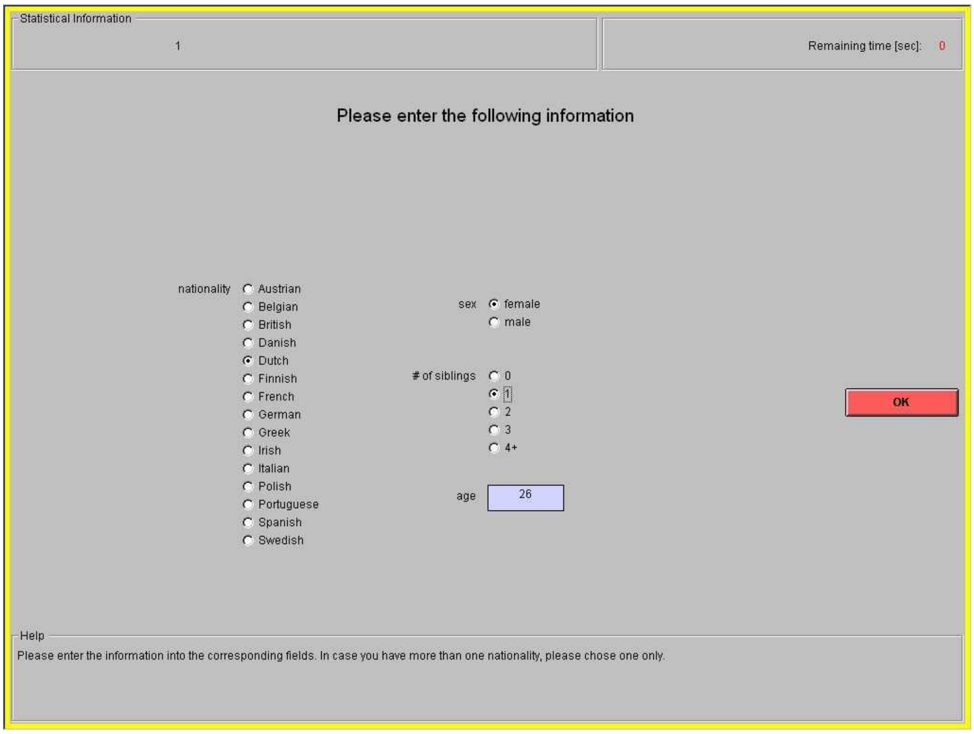


Figure 6: Screenshot of the CV information

They also had a hard copy of the instructions next to their machines.[11] The instructions were followed by a short quiz of three questions covering the crucial aspects of the game. Almost all participants appeared to have understood the game very well before playing. No major clarification questions were asked. After reading through the instructions, participants were asked to enter information about their age, gender, nationality, and number of siblings.[12] To increase anonymity, the age displayed to fellow players was modified by adding a random number. This was also mentioned in the instructions further to a general anonymity and privacy statement.

Each session consisted of six repetitions in which participants were randomly matched in groups of five players. In this experiment we only use the first four repetitions.

In each of these four repetitions participants played the following repeated version of the trust game. Figures 6 to 9 provide examples of the relevant screenshots seen by participants. At the beginning of the repetition, each player could see some information about the four other players in the own group, the information included the players' nationality, age, gender, and the number of siblings. The participants then decided to *whom* and *how much* of their initial endowment of 100 they were willing to transfer. No entry in any of the boxes corresponded to making no choice, which was also an option. In the next step participants saw *who* among the other players had chosen them and

[11] At `http://www.kirchkamp.de/pdf/trustInstructions.pdf` you can download a copy of the instructions

[12] During the recruitment process it was made sure that participants were recruited only from countries which have a substantial number of students at the university. This restriction was introduced to avoid identification of the participants during the game.

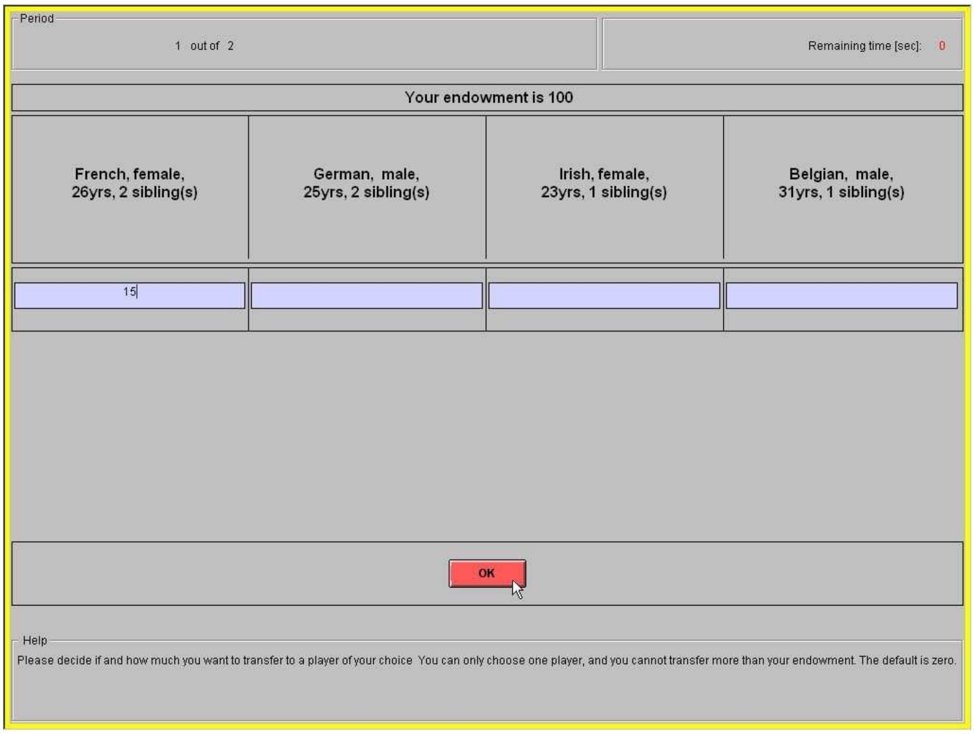


Figure 7: Screenshot of the first stage

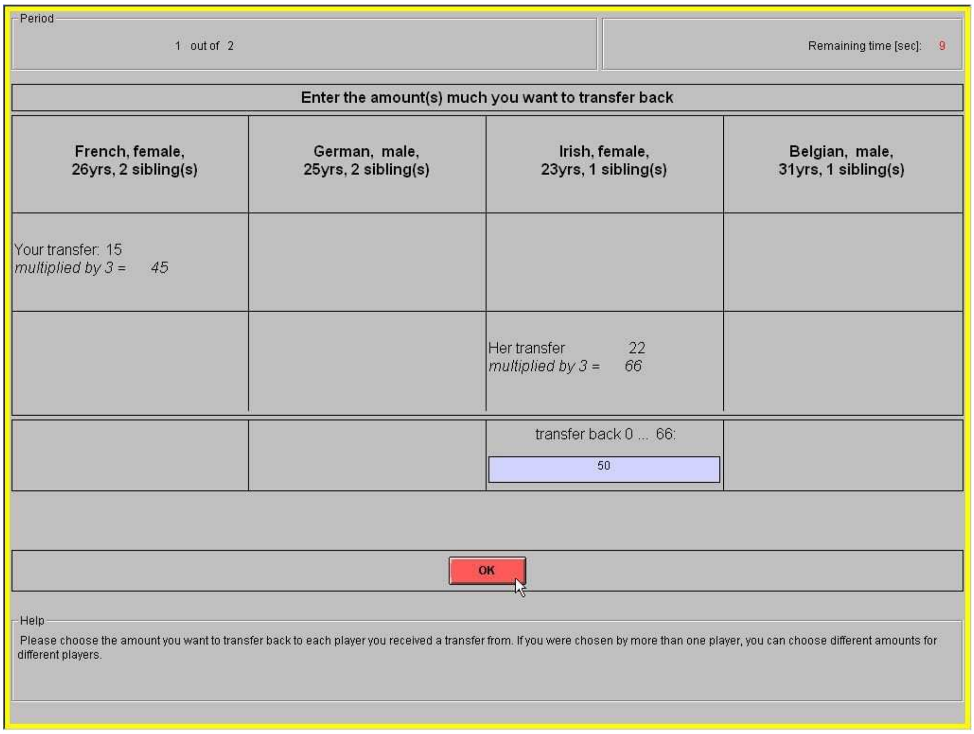


Figure 8: Screenshot of the second stage

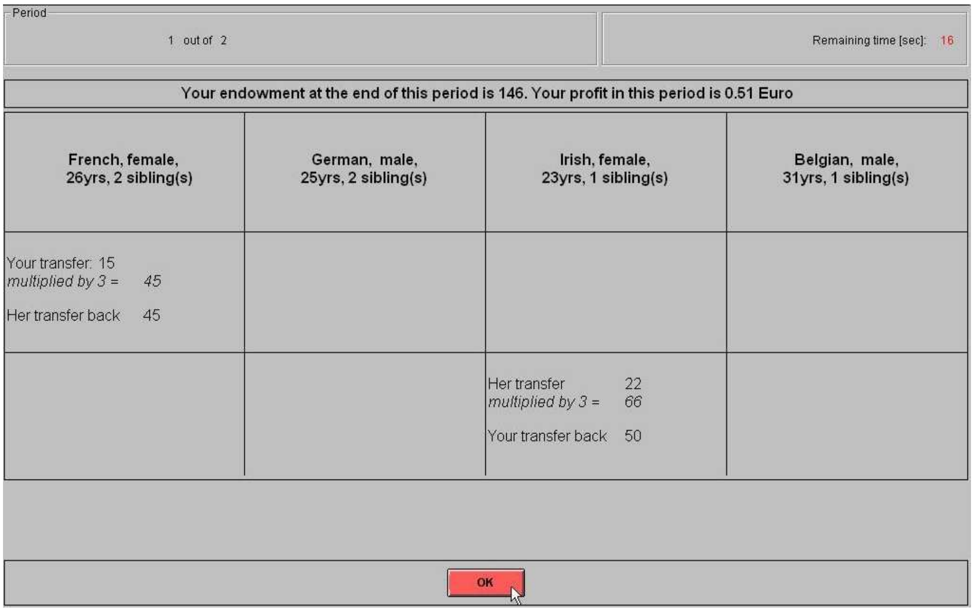


Figure 9: Screenshot of the third stage

how much they had received from these partners. In addition, this amount was shown multiplied by three. For each player from whom a transfer was received, they could choose how much to return. Then, participants were presented a summary of all transfers and returns they had been involved with. These steps were repeated 6 times. Then, groups were reshuffled and a new repetition was played. Due to the limited amount of participants in each session and the large size of each group, the re-matching had to be done on a random basis, hence it is not ruled out that participants could meet again in subsequent groups. At the end of each repetition participants were also informed about their own profit made over all the periods of that repetition.

B Regression Results

All regressions are based on models with mixed effects. Standard errors, t-statistics, p-values, and confidence intervals are based on a parametric bootstrap based in 1000 replications.

B.1 Transfers and Returned Amounts

We assess the significance of our discussion of Figure 1 with the help of two mixed effects regressions.

In the first one we investigate how trust as measured by the amount of tokens sent depends on other variables. The amount of tokens sent is denoted by t^S_{ij} where i refers to the identity of the participant and j is a *period number* that uniquely identifies the period and the repetition, so $j = 1, .., 24$. In a second regression we investigate how the amount of tokens returned depends on the sum of transfers received and on other variables. The amount of tokens returned by participant i to all those from whom tokens were received

	β	σ	t	p value	95% conf	interval
d_1	54	3.8	14.2	0.0000	46.5	61.4
d_{2-5}	65	3.71	17.5	0.0000	57.7	72.3
d_6	43.1	3.76	11.5	0.0000	35.7	50.5
T	5.96	0.425	14	0.0000	5.13	6.79

Table 10: Estimation of equation (7), transfer t^{S}

in that period is denoted by r^S_{ij} where j is the index of the period of a given repetition. The sum of transfers received by participant i in the period numbered j is denoted by t^R_{ij}. So t^R_{ij} is the tripled amount of tokens sent by other players to participant i in the period numbered j. The coefficient of t^R_{ij} in the regression can be considered as a measure of marginal trustworthiness. Specifically we run the following regressions:

$$\begin{aligned} t^S &= \beta_{d_1} \cdot d_1 + \beta_{d_{2-5}} \cdot d_{2-5} + \beta_{d_6} \cdot d_6 + \beta_T \cdot T + \epsilon_s + \epsilon_i + \epsilon_{ij} && (7) \\ r^S &= \beta_{d_1} \cdot d_1 + \beta_{d_{2-5}} \cdot d_{2-5} + \beta_{d_6} \cdot d_6 + \beta_T \cdot T \\ &\quad + \beta_{t^R_1} \cdot t^R_1 + \beta_{t^R_{2-5}} \cdot t^R_{2-5} + \beta_{t^R_6} \cdot t^R_6 + \beta_{t^R_T} \cdot t^R_T + \epsilon_s + \epsilon_i + \epsilon_{ij} && (8) \end{aligned}$$

Sessions are indexed with s, participants are indexed with i, and j is the period number. To simplify notation we do not write indices $_{ij}$ for variables. Throughout the paper and unless specified otherwise we estimate mixed effect models with random effects ϵ_s, ϵ_i and ϵ_{ij} for session s, participant i, and participant i in period j, respectively. We assume that error terms ϵ_s, ϵ_i and ϵ_{ij} are independent and follow a normal distribution with mean zero. With this specification, we allow behavior of the same participant in different repetitions to be correlated as well as different participants from the same session to be correlated. This is important as participants within the same session are randomly assigned to groups in each repetition and thereby potentially influence each other. Note that each participant belongs to a unique session and, hence, session indices are only needed in the error terms. Dummies d_1, d_{2-5} and d_6 are one in period 1, periods 2-5 and period 6, respectively, and zero otherwise. We let t^R_1, t^R_{2-5} and t^R_6 specify the transfer received in period 1, periods 2-5 and period 6, respectively, so t^R_k is short for $t^R \cdot d_k$ for $k \in \{\text{“1”}, \text{“2-5”}, \text{“6”}\}$. Since behavior might change over time we include the repetition $T \in \{1, 2, 3, 4\}$ of the experiment. Results are presented in Tables 10 and 11.

The Tables 10 and 11 confirm what we see in Figure 1. Trust increases during the initial stage of a repetition (β_{d_1} is significantly smaller than $\beta_{d_{2-5}}$ in equation (7), $p < 0.0001$) and decreases at the end ($\beta_{d_{2-5}}$ is significantly larger than β_{d_6}, $p < 0.0001$). In fact, trust in the final period is lower than in the first period (β_{d_1} is significantly larger than β_{d_6}, $p < 0.0001$). Trust increases during the experiment (β_T is significantly positive).

Consider now trustworthiness. The sum of total returns r^S reacts mainly to the sum of transfers t^R received. Marginal trustworthiness, measured as coefficients of t^R_1, t^R_{2-5} and t^R_6, is significantly positive during all periods. Neither β_T nor the intercepts β_{d_1}, $\beta_{d_{2-5}}$, β_{d_6} are significant.

The coefficients $\hat{\beta}_{t^R_1} = 0.515$ and $\hat{\beta}_{t^R_{2-5}} = 0.541$ capture the estimated marginal trustworthiness in periods 1 and 2-5, respectively. Both coefficients are significantly above

	β	σ	t	p value	95% conf	interval
d_1	-0.00893	11.8	-0.000757	0.9994	-23.1	23.1
d_{2-5}	-0.899	10.9	-0.0827	0.9341	-22.2	20.4
d_6	-3.14	11.6	-0.27	0.7871	-25.9	19.6
T	2.41	1.86	1.29	0.1962	-1.24	6.05
t_1^{R}	0.515	0.0242	21.3	0.0000	0.467	0.562
t_{2-5}^{R}	0.541	0.0181	29.8	0.0000	0.506	0.577
t_6^{R}	0.328	0.0242	13.6	0.0000	0.281	0.376
t_T^{R}	0.0136	0.00613	2.22	0.0262	0.00161	0.0257

Table 11: Estimation of equation (8), returned amount r^{S}, this is the sum of potentially several amounts

1/3 ($p < 0.0001$).[13] Thus, we find strong evidence that participants are trustworthy (at the margin). There is no significant evidence that trustworthiness differs in the period 1 from periods 2-5 ($p = 0.1592$) but we do find a significant endgame effect. Trustworthiness decreases in the last period to $\hat{\beta}_{t_6^R} = 0.328$ which is significantly different from $\beta_{t_{2-5}^R}$ ($p < 0.0001$). Trustworthiness in the final period is so low that there is no longer significant evidence, as in periods 1-5, that senders get back more than they sent ($\beta_{t_6^R}$ is not significantly different from 1/3, $p = 0.8450$). Finally, note that trustworthiness increases significantly between repetitions as $\beta_{t_T^R}$ is significantly positive.

B.2 Robustness of the Results Presented in Section 3.3

In section 3.3 we discuss determinants of success. We find that participants from the North earn significantly more in our experiment than participants from the south. How much do these results depend on our categorization. To check this we present the following two exercises. First we drop d_N in (1) and include instead latitude ϕ and longitude λ of the respective countries from the CIA database. This leads to the following regression:

$$\begin{aligned} \pi_i \;=\; & \beta_1 + \beta_\phi \cdot \phi + \beta_\lambda \cdot \lambda + \beta_{d_{♂}} \cdot d_{♂} + \beta_A \cdot A \\ & + \beta_{d_{S_1}} \cdot d_{S_1} + \beta_{d_{S_2}} \cdot d_{S_2} + \epsilon_s + \epsilon_i \end{aligned} \qquad (9)$$

Results are shown in Table 12.

While latitude might be a rather naïve predictor for success it is still the only significant coefficient. Note that it is also the coefficient with the largest pmvd value of 0.847. It is remarkable that latitude as a very crude measure of difference between participants captures some differences in success while longitude does not. This regression remains only a robustness check as we do not expect that latitude matters per se but instead that the cultural similarities among the countries further north and further south could play a role. The mean latitude remains an arbitrary albeit focal divide between these regions. As a second exercise we do an approximate permutation test, i.e. we estimate (1) again,

[13] Recall that tokens received is equal to the tripled amount of tokens sent. So if $\beta_{t_{2-5}^R} \geq 1/3$ then the sender gets back more than she sent if she decides to send marginally more tokens.

	β	σ	t	p value	95% conf	interval	pmvd
1	115	55.5	2.07	0.0407	4.96	225	
ϕ	2.05	0.775	2.64	0.0096	0.511	3.59	0.847
λ	0.125	0.591	0.211	0.8336	-1.05	1.3	0.005
$d_{♂}$	-3.7	10	-0.369	0.7128	-23.6	16.2	0.014
A	1.01	1.7	0.593	0.5542	-2.36	4.37	0.050
d_{S_1}	8.2	15.2	0.538	0.5919	-22	38.4	0.023
d_{S_2}	13.8	16.3	0.852	0.3965	-18.4	46.1	0.061

Table 12: Determinants of success — estimation of equation (9) payoff π

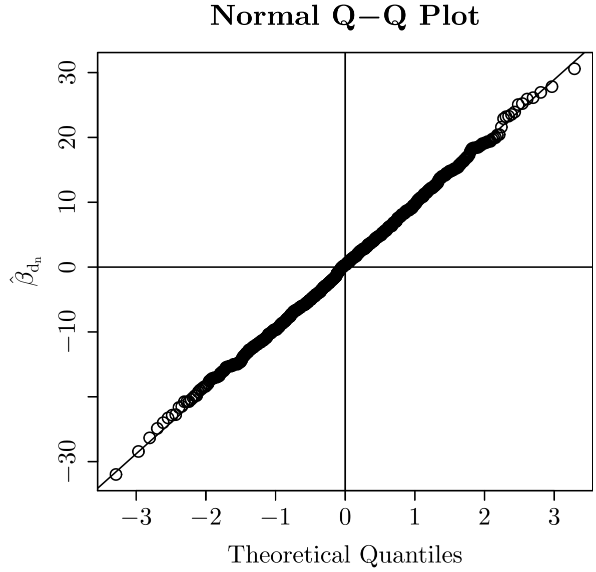


Figure 10: Q-Q plot for 1000 estimates of β_{d_N} of equation (1) with a random d_N dummy

but replace d_N with a random dummy. This dummy has the same number of zeros and ones as the original dummy d_N.

We estimate equation (1) 1000 times, each time with a new random d_N dummy. Each time we get a different estimate for β_{d_N}. A Q-Q plot for $\hat{\beta}_{d_N}$ is shown in figure 10. We find that it is rather unlikely ($p = 0.0040$) to accidentally get an estimate for β_{d_N} that is greater than the value of 25.5 determined by the data and indicated in Table 1.

B.3 Payoff Comparison North and South, per Period

To support our discussion of figure 5 in section 3.4, and to come to point 4 from the introduction, we run the following regression

$$\pi_i = \beta_1 + \beta_{d_N} \cdot d_N + \beta_P \cdot P + \beta_{PN} \cdot d_N \cdot P + \beta_6 \cdot d_6 + \epsilon_s + \epsilon_i \,. \qquad (10)$$

In this estimation we capture the trend of South in β_P, the trend of North in β_{PN} and control for a constant effect in the last period by adding d_6. Results are shown in Table

	β	σ	t	p value	95% conf	interval	pmvd
1	238	11.5	20.6	0.0000	215	260	
d_N	4.27	14.4	0.297	0.7662	-23.9	32.5	0.007
P	2.09	2.68	0.779	0.4361	-3.16	7.33	0.012
$P \cdot d_N$	6.27	3.21	1.95	0.0507	-0.02	12.6	0.411
d_6	-54.8	9.54	-5.75	0.0000	-73.5	-36.1	0.570

Table 13: Determinants of success — estimation of equation (10) payoff π

13.

While we find no significant change in the success of South we observe a significant increase in success of North across periods. The average drop in success in period 6 is strongly significant and substantial (estimated to be -54.8 tokens).

B.4 Details of the Calcuations in Section 3.5

The estimated increase in marginal trustworthiness among North is $\beta_{t^R_{TN}} + \beta_{t^R_T} = 0.0467 + (-0.0109) = 0.0358, \mathrm{CI}_{95} = [0.0193, 0.0522]$

The average transfer in period 1 is $54 + 2.5 \cdot 5.96 = 68.9$ where the multiplier 2.5 is the average number of repetitions.

North returns on average $(12.2 + (-5.97) + 2.5 \cdot (5.74 + (-6.71)) + (54 + 2.5 \cdot 5.96) \cdot 3 \cdot ((0.543 + (-0.0517)) + 2.5 \cdot (0.0467 + (-0.0109)))) = 124, \mathrm{CI}_{95} = [96.4, 151]$ in period 1. Hence, the return ratio is $(12.2+(-5.97)+2.5\cdot(5.74+(-6.71))+(54+2.5\cdot5.96)\cdot3\cdot((0.543+(-0.0517)) + 2.5 \cdot (0.0467 + (-0.0109))))/3/(54 + 2.5 \cdot 5.96) = 0.599, \mathrm{CI}_{95} = [0.48, 0.717]$. Similarly we estimate the return ratio for South in period 1 to be $((-5.97) + 2.5 \cdot 5.74 + (54+2.5\cdot5.96)\cdot3\cdot(0.543+2.5\cdot((-0.0109))))/3/(54+2.5\cdot5.96) = 0.556, \mathrm{CI}_{95} = [0.432, 0.68]$. For periods 2-5 we find analogously an estimated return ratio of North and of South equal to $0.601, \mathrm{CI}_{95} = [0.502, 0.699]$ and $0.592, \mathrm{CI}_{95} = [0.493, 0.69]$, respectively. For period 6 we find $0.351, \mathrm{CI}_{95} = [0.203, 0.499]$ and $0.408, \mathrm{CI}_{95} = [0.262, 0.555]$, respectively.

B.5 Details of the Calcuations in Section 3.6

Based on Table 2 we determine that North sends on average $8.84 + 2.5 \cdot (-0.274) = 8.15$ more tokens than South in periods 2-5 (The number 2.5 is again the average number of repetitions). This means that those selected by North obtain $(8.84+2.5\cdot(-0.274))\cdot3 = 24.5$ more tokens. Following Table 11, $24.5 \cdot 0.544 = 13.4$ tokens are returned. Hence, the difference in transfer between North and South in periods 2-5 generates a net gain of $(8.84 + 2.5 \cdot (-0.274)) \cdot (3 \cdot 0.541 - 1) = 5.08, \mathrm{CI}_{95} = [0.707, 9.46]$ more tokens for North.

Following Table 3 we observe that North returns initially 16.2 tokens more than South, but at the margin (for each token received) -0.105 fewer tokens than South. We take the estimate of the average transfer in periods 2-5 from Table 10 which is 65 and, thus, estimate that in periods 2-5 North earn due to the difference in their return behavior $-(-0.105) \cdot 65 \cdot 3 - 16.2 = 4.35, \mathrm{CI}_{95} = [-16.4, 25.1]$ more tokens on average than South. The combined effect for periods 2-5 is, thus, $5.08 + 4.35 = 9.43, \mathrm{CI}_{95} = [-11.8, 30.6]$. Similarly, we calculate the effect for period 1 as $3.21, \mathrm{CI}_{95} = [-21.1, 27.5]$ and for period

6 as $6.05, \text{CI}_{95} = [-19.6, 31.7]$. The average effect for all 6 periods is, hence, $7.83, \text{CI}_{95} = [-13, 28.7]$ which is considerably, although not significantly, below the estimated difference in payoffs of 25.5 in Table 1.

C Conditional Logit Model

The estimates of Table 6 are computed using the conditional logit model proposed by McFadden (1973). This model is applied to our setting in the following way. Define U_{ikp} as the utility of i if i chooses k in period p, and

$$d_{ikp} = \begin{cases} 1 & \text{if } i \text{ chooses } k \text{ in } p, \\ 0 & \text{otherwise.} \end{cases}$$

The time index p stands for the six periods of the game. Conditional on participating in the game (*i.e.* not making a zero transfer), each player can choose one among four possible partners, $k = \{1, 2, 3, 4\}$ in each period. The four choices are mutually exclusive and exhaustive. The random utility corresponding to each choice is assumed to be:

$$U_{ikp} = \alpha d_{ikp-1} + \delta d_{kip-1} + \nu X_{ik} + \epsilon_{ikp}$$

for $k = \{1, 2, 3, 4\}$. Using the above notation, d_{ikp-1} means that player i has chosen k in the previous period. Similarly, d_{kip-1} means that player k has chosen player i in the previous period. The omitted variable is that player i and k had no interaction in the previous period. The other covariates X_{ik} include the remaining choice specific characteristics such as gender, nationality (both interacted with the corresponding attributes of i), age, and siblings. Note that the fact that k was previously chosen by i (or not) is interpreted as a characteristic of the choice k in p. By the same token, the fact that i was chosen by k (or not) in period $p - 1$ becomes a characteristic of k in p. Hence, previous playing behavior can be seen as generating observable choice specific attributes in p.

Player i chooses player k if this yields highest utility. Hence,

$$\Pr(d_{ikp} = 1) = \Pr(U_{ikp} > U_{ilp}) : \forall : l \neq k.$$

The estimates from this model are reported in column 1 of Table 6. All estimated coefficients are reported in the form of odds ratios.

In column 2 of Table 6 we add dummies that account for whether North has chosen a participant from North and similarly whether South has chosen someone from South.

This random utility model can be augmented by adding variables which characterize the effect of previous behavior in more detail. In column 3 of Table 6 we interact the dummy indicating whether i transferred to k in the previous period with a dummy indicating whether k returned more than the median return ratio in the sample. Similarly, we interact the dummy indicating whether k transferred to i in the previous period with a dummy indicating whether k transferred more than the median transfer in the sample.

Tables 7 and 9 redo the previous analysis to the sample restricted to choices in periods 2 and 6, respectively. Table 9 includes interactions with whether the participant belonged to region North.

Table 14: Nationalities: frequencies and average latitude

country	av. latitude	participants
Southern countries		
Greece	39	9
Portugal	39.3	1
Spain	40	11
Italy	42.5	17
France	46	12
Austria	47.2	6
Northern countries		
Belgium	50.5	5
Germany	51	16
Poland	52	3
Netherlands	52.3	8
Ireland	53	5
United Kingdom	54	8
Denmark	56	3
Sweden	62	4
Finland	64	2

Source: CIA (2003).