

UNIVERSITÀ DEGLI STUDI DI SALERNO

Dipartimento di Studi Economici e Statistici

Tesi di Dottorato in Economia del Settore Pubblico
Corso di Dottorato in Analisi Economica, Giuridica e Statistica delle Politiche,
dei Mercati e delle Imprese (XXIX ciclo)

State-contingent Trust and Trustworthiness
An experimental analysis

Coordinatore

Tutor

Prof. Sergio Destefanis

Prof. Marcello D'Amato



Candidato

Marco Stimolo

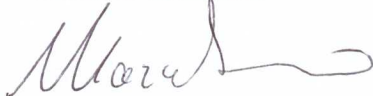


Table of Contents

| | |
|---|-----------|
| Chapter 1: Introduction | 6 |
| 1.1 Definition of the topic | 6 |
| 1.2 Experiments on trust and trustworthiness | 7 |
| 1.3 State-contingent trust and trustworthiness | 10 |
| 1.3.1 Experimental investigations | 10 |
| 1.3.2 Overview of the main results | 13 |
| 1.4 Structure of the thesis | 15 |
| 1.5 References | 15 |
| CHAPTER 2: Experimental Investigations on Trust | 20 |
| 2.1 Introduction | 20 |
| 2.2 What are trust and trustworthiness? | 24 |
| 2.2.1 Trust and trustworthiness as rational behaviors | 24 |
| 2.2.2 Trust and trustworthiness as cultural traits | 25 |
| 2.2.3 Trust and Trustworthiness as evolutionary traits | 28 |
| 2.2.4 Young people's trust and its consequences for the labor market | 30 |
| 2.3 The trust game | 32 |
| 2.4 Determinants of trust and trustworthiness | 34 |
| 2.4.1 Trust and trustworthiness in the lab | 34 |
| 2.4.2 Trust and trustworthiness are not pure altruism | 36 |
| 2.4.3 Distinguishing trust and trustworthiness from unconditional other-regarding preferences | 38 |
| 2.4.4 Trusting behavior and individuals' attitude towards risk | 41 |
| 2.5 Correlation between survey and experimental data on trust and trustworthiness | 44 |
| 2.6 State-contingent trust and trustworthiness | 47 |
| 2.7 Conclusions | 50 |
| 2.8 References | 51 |

CHAPTER 3: Trust and reciprocity in youth labor markets. An experimental approach to analyzing the impact of labour market experiences on young people 56

| | |
|--|----|
| 3.1 Introduction | 56 |
| 3.2 Experimental Design | 59 |
| 3.3 Results | 64 |
| 3.3.1 Senders' behavior in TG1 (no information on the counterpart) | 64 |
| 3.3.1.1 Trust game 1: Econometric analysis of trusting behavior | 68 |
| 3.3.2 Trust Game 1: Receivers' behavior | 71 |
| 3.3.2.1 Trust Game 1: Econometric analysis of trustworthiness | 74 |
| 3.3.4 Trust Game 2: The impact of information on recipients | 76 |
| 3.3.4.1 Econometric analysis of changes in trusting behavior | 81 |
| 3.3.5.1 Econometric analysis of changes in trustworthiness | 86 |
| 3.4 Conclusions | 87 |
| 3.5 References | 89 |

CHAPTER 4: State-contingent trust and trustworthiness. An experimental investigation 94

| | |
|--|-----|
| 4.1 Introduction | 94 |
| 4.2 Experimental design | 98 |
| 4.3 Results | 101 |
| 4.3.1 Dictator Game | 101 |
| 4.3.2 Differences in senders' behavior in Dictator game 1 and Trust game 1 | 105 |
| 4.3.2.1 Receivers' behavior in trust game 1 | 108 |
| 4.3.3 Differences in senders' behavior between trust game 1 and trust game 2 | 109 |
| 4.3.3.1 Difference in senders' behavior between TG1 and TG2 - econometric analysis | 115 |
| 4.3.4 Differences in receivers' behavior between trust game 1 and trust game 2 | 117 |
| 4.3.4.1 Difference in receivers' behavior between TG1 and TG2 - Econometric analysis | 123 |

| | |
|--|------------|
| 4.3.5 Differences in senders' behavior between trust game 2 and trust game 3 (information on players' state) | 124 |
| 4.3.5.1 Differences in senders' behavior between TG2 and TG3 - Econometric analysis | 127 |
| 4.3.6 Differences in receivers' behavior between trust game 2 and trust game 3 (information on players' state) | 128 |
| 4.3.6.1 Differences in receivers' behavior between Trust game 2 and Trust game 3 - Econometric analysis | 133 |
| 4.4 Conclusions | 135 |
| 4.5 References | 137 |
| General Conclusions | 140 |

Chapter 1: Introduction

1.1 Definition of the topic

This thesis investigates experimentally how the individuals' state (e.g. economic, labor and social status) influences their trusting and trustworthy behavior in one-shot anonymous interactions. Moreover, it analyzes the way the knowledge of the counterpart's state mediates subjects' behavioral trust and trustworthiness.

Trusting and trustworthy behavior arise in repeated (anonymous or non-anonymous) social interactions where the truster exposes himself to the trustee's discretionary power to either lower or increase his utility (Hardin 2002). Economists generally rely on Diego Gambetta's definitions of trust and trustworthiness: "when we say that [...] someone is trustworthy, we implicitly mean that the probability that he will perform an action that is beneficial [...] is high enough for us to consider to engaging in some form of cooperation with him" (Gambetta 2000: 216). On this account, trust and trustworthiness stem from individual rationality and self-interest. Indeed, the trustee may have a interest in repaying trust in repeated anonymous or non-anonymous interactions to better fulfill his self-interest in the long-run (Hardin 2002, Gauthier 1986). Hence, a trusting choice derives from the truster's selfish preferences and from his (*rational*) belief on the trustee: i.e. a person chooses to entrust someone if she assesses the expected return to trusting to be higher than the expected return to not trusting.

In one-shot anonymous interactions standard theory predicts trust and trustworthiness not to emerge because they are not sustainable as a Nash-equilibrium outcome. On this account, the trustee does not have any incentive to repay trust; anticipating this, the truster rationally chooses not to entrust the counterpart. Nonetheless, the dynamics of impersonal trade in large markets consist of one-shot anonymous interactions, whose efficient outcome strictly depends on individuals' trust and trustwor-

thiness. Indeed, a recent surge of empirical survey-based research provides evidence of a significant correlation between the average level of trust in a country and important economic variables such as economic development (Dearmon and Grier 2009), GDP (Algan and Cahuk 2010), inflation and the volume of trade between countries (Guiso et al. 2009). Hence, the contradiction between the Nash-equilibrium prediction and the evidence of trust and trustworthiness in impersonal trades calls for the identification of the motivational determinants of these behavioral regularities. Next section illustrates the experimental method widely used to tackle this issue.

1.2 Experiments on trust and trustworthiness

The survey-based analyses mentioned above exhibit several limits for the the identification of the motivations inspiring trusting and trustworthy behaviors. Specifically, they cannot validate the behavioral implications of the attitudinal questions about trust and trustworthiness. Given the absence of any explicit incentive to report truthfully, subjects can be prone to social desirability bias (i.e. subjects tend to answer in a way that they believe is socially desirable). Thus, the consistency of self-reported trust and trustworthiness with trusting and trustworthy behavior is not granted. This strongly limits survey-based analysis for the identification of the motivations inspiring trust and trustworthiness. Because of that, the last twenty years have seen the progressive development of experimental analyses of behavioral trust and trustworthiness.

These investigations focus on the trust game (Berg et al. 1995), which involves a sender (the truster) and a receiver (the trustee), both endowed with a monetary amount. The sender chooses whether to give some or all of it to the receiver. The latter gets (typically) three times the amount originally sent and chooses whether to give back some or all of it. Standard theory is based on the auxiliary assumption of self-interest and predicts a zero level of net transfers: i.e. a rational sender knows that a rational

receiver will not send anything back and so there is no (financial) benefit from sending anything. However, senders do trust and the receivers do reciprocate; a recent meta-analysis of trust games (Johnson & Mislin, 2011) has found that, on average, senders give around 50% of their allocation, whilst receivers send back 40% of the amount received.

The trust game represents the logic of different kinds of social interactions, from economic exchanges (e.g. buyer-consumer interactions, financial markets and principal-agent interactions) to personal interactions where the objects of trade are “gifts, assistance and favors” (Smith 1998: 3). Because of that, a trusting behavior and the decision to repay trust are compatible with other-regarding motivations, as both the truster and the trustee may have an intrinsic interest in the well-being of the counterpart, so to benefit her independently of any future return. To elicit individuals’ altruism and identify its contribution to the explanation of trusting and trustworthy behavior (Cox 2004, Ashraf *et al.* 2006) the Dictator Game has been added to the trust game. In the dictator game, only the sender is endowed with a monetary amount and simply chooses whether to give some or all of it to the recipient, who is passive and cannot reciprocate. The evidence of altruism in the dictator game is used as a base line with any increase in the amount sent by the sender (or sent back by the receiver) in the trust game being imputed to trust or reciprocity (Cox 2004).

The violation of the assumption of selfish preferences entails that the players of the trust game can be of different types: selfish, altruistic and reciprocal. This heterogeneity in players’ motivations to trust and trustworthiness introduces *uncertainty* about the payoff functions of the other players and it legitimates different beliefs about their behavior. This uncertainty entails that individuals’ risk attitude can influence their level of trust. Specifically, for any given belief about others’ trustworthiness, the trusters (i.e. senders) may differ in their propensity to take the risk to be cheated on by the trustees (i.e. receivers). Thus, individual decision problems under risk (i.e. lottery choice) are generally used to elicit subjects’ level of risk aversion to

test the hypothesis that for a given level of beliefs about others' trustworthiness, the more risk averse a person is, the lower her level of trust will be (Houser et al. 2009).

Since individuals learn to trust by repeatedly engaging in anonymous or non-anonymous interactions in a given context, their past experiences and socio-economic characteristics are likely to determine their heterogeneity in trusting and trustworthy behavior in anonymous one-shot interactions. Since these characteristics of individuals' preferences and beliefs are not inferable from behavioral data, recent analyses combine experimental investigations with surveys-based measures of individuals' attitudes towards trust and trustworthiness. This approach uses the experimental method to validate the behavioural content of subjects' answers to survey questions.

Amongst these contributions, Fehr et al. (2003)'s experiment identifies a negative effect of unemployment on the level of receivers' transfers. This result raises the issue of a dependency of behavioral trust and trustworthiness from individuals' state: i.e. the levels of trust and trustworthiness are likely to vary in accordance with their economic, labor or social status. Fehr et al. (2003)'s study is, however, purely correlational as the variable inherent to subjects labor market status is not exogenously determined. To identify the causal link between subjects' state and their trusting/trustworthy behavior, Anderson et al (2006) and Hargreaves Heap et al. (2013) provide an experimental design based on the exogenous variation of subjects' endowment. Their experimental results, however, lack robustness and an analysis of the interaction of induced heterogeneity in state with the multiple motivations (i.e. selfish or altruistic) to trust and trustworthiness is still lacking.

The next section illustrates the way this thesis contributes to the study of state-contingent trust and trustworthiness, with particular regard to the interaction between individuals' state and the different motivations to trust and trustworthiness.

1.3 State-contingent trust and trustworthiness

To improve the experimental analysis of state-contingent trust/trustworthiness, this thesis reports the results of two experiments aimed at identifying the systematic relationship between subjects' state and their trusting/trustworthy behavior as well as the effect of the knowledge of players' state on their behavioral trust and trustworthiness. This analysis is explorative in character because of two main reasons. First, as clarified in the previous section, the results on the effect of exogenously induced heterogeneity in state on trust and trustworthy behavior are not robust; hence, the existence of a causal nexus cannot be taken for granted at the outset. Second, no theoretical priors are available to predict the sign of this causal link.

1.3.1 *Experimental investigations*

Both experiments are based on a within-subjects design, where the same person participates to different behavioral tasks. This method allows to control for the high variance in treatment effect due to randomization by essentially eliminating the impact of the subject-specific effect (List et al. 2011). Consistently with this methodological choice, the analysis is mainly based on the within-subjects behavioral variations across the different behavioral tasks. In addition to that, the experimental designs allows to compare the behavioral variations of different samples of subjects.

The first experiment identifies the systematic differences in trusting and trustworthy behavior between individuals in different labor market states - i.e. employed, student and NEET (Not in the Employment, Education or Training)¹ - as well as the ef-

¹ Of course the term 'NEET' was not in broad use until well into the new millennium when the OECD began to employ it. Previously, 'joblessness', 'non-employment', 'broad unemployment', 'discouraged workers' were all terms which found gainful usage – often with different meanings being applied by different authors – to describe the NEET in part or in their entirety.

fect of the information on the counterpart's state on subjects' behavioral trust and trustworthiness.

Trust and trustworthiness allows people to engage in mutually beneficial transactions. Hence, more trusting or trustworthy individuals have more chances to be successful in the job search. Moreover, trust is positively correlated to job stability as individuals with a higher propensity to trust and trustworthiness are more cooperative in the job place, so to increase firms benefits and - consequently - reducing the probability of being fired. The opposite argument holds for less trusting or trustworthy individuals. The success in the job search and the status of employed further boost individuals' willingness to trust others or to be trustworthy. Hence, employed people should exhibit a higher level of trust and trustworthiness with respect to NEETs. This state-contingency of trust and trustworthiness entails that the information on the labor market status of the counterpart may influence subjects' trusting/trustworthy behavior.

The core of the first experimental design is based on two one-shot trust games with random and anonymous matching played in different informational conditions. In the first one, subjects in different labor market states play the game in the standard way. In the second trust game, subjects are randomly and anonymously re-matched and their labor market state is made public. In line with the contribution of Fehr et al. (2003), this experiment tries to assess whether subjects' labor market status elicited through survey questions is predictive of their trusting and trustworthy behavior in the lab. Specifically, this experiment focuses on the internal heterogeneity of the categories of employed and unemployed, by analyzing the differences in behavioral trust and trustworthiness between permanent or precarious employed as well as between NEETs unemployed (i.e. peoples searching for a job without success) and other NEETs (people out of labour forces and not active in the job search). In addition to that, this experimental investigation aims at investigating the way subjects' altruism and risk attitude interact with their labor market status in influencing their

level of trust and trustworthiness. Hence, dictator games and lottery choice problems are implemented to control for individuals' altruism and risk aversion.

The first experiment, however, cannot identify the causal nexus between individuals' labor market status and their trusting/trustworthy behavior because of a self-selection problem. Indeed, individuals' can self-select in advantageous or disadvantageous job positions because of their preferences' characteristics (e.g. altruism and self-interest) or because of biased beliefs about others' trustworthiness inhibiting their willingness to cooperate. This entails a problem of reverse causality as the proposed experiment cannot identify the right direction of the causal link between the labour market state and individuals' behaviour; are people unemployed because they do not entrust others or *vice versa*? The first experiment cannot answer this question.

To address this issue, the second experiment identifies the effect of the exogenously induced heterogeneity in state on behavioral trust and trustworthiness, as well as the behavioral effect of the public knowledge of the counterpart's state. The experiment is based on three trust games with random and anonymous matching played in different states and informational conditions. In the first one, subjects get an *equal* endowment and play the game in the standard way. In the second one, a random mechanism induces a distribution of "Disadvantaged", "Median" and "Advantaged" players by varying their endowment. In the third trust game, subjects maintain the state determined in the second one, but they get informed on the counterpart's state. This treatment allows understanding the way the information on the counterpart's state mediates subjects' behavioral trust and trustworthiness. To analyze the interaction between the exogenous variation of subjects' state and the different motivations to trust/trustworthiness, the experimental protocol implements dictator games and lottery choice problems to control for altruism and risk aversion.

1.3.2 Overview of the main results

The first experiment shows that the categories of NEET and employed exhibit an internal heterogeneity at the behavioral level. Specifically, the sub-category of NEETs unemployed (i.e. NEET searching for a job *without success*) are the least trustworthy. By contrast, the other categories of NEETs (NEETs not searching for a job) and employed (specifically those with a *permanent* contract) are relatively more trusting and trustworthy. This entails that the experience of failure in the job search mediates the detrimental effect of unemployment on behavioral trust and trustworthiness. Furthermore, the sub-category of precarious workers (i.e. fixed term or informal workers) tend to exhibit lower levels of behavioral trust and trustworthiness. Hence, the increasing flexibility of labour contracts is likely to have long lasting consequences on young people labor market by reducing their level of trust and trustworthiness. Unemployed NEETs are more disadvantaged than other sub-categories of NEETs that do not have the need to be active in the job search. Precarious workers are likely to be similarly disadvantaged not only in terms of income, but also in terms of experience in the job market. Thus, the first experiment shows that the most disadvantaged categories of subjects systematically exhibit lower levels of behavioral trust and trustworthiness.

The analysis of the second trust game (information on the counterpart's labor market status) clearly shows that NEETs and Employed tend to cooperate less between each other than with other categories of players. This result, together with the evidence of lower levels of trustworthiness from unemployed NEETs, suggests a candidate mechanism for the micro-level explanation of the persistence of long-term wage and employment penalties: The systematic negative correlation of the unemployed state with young individuals' reliability provides the rationale for employed people to di-

trust them; this may prevent a successful search for a job in the long run to young unemployed people.

The results of the second experiment show that the exogenously induced heterogeneity in state significantly reduces the average level of trust from “Disadvantaged” senders. Similarly, induced heterogeneity reduces the average level of trustworthiness of all the categories of receivers, particularly of “Disadvantaged” ones. By contrast, “Advantaged” players do not exhibit a clear pattern of trusting and trustworthy behavior. Thus, “Disadvantaged” players behave like unemployed NEETs in the first experiment, supporting the hypothesis of a negative causal effect of the unemployed state over trusting and trustworthy behavior.

In the third trust games (information on the counterpart’s state) “Disadvantaged” subjects *increase* their level of trust and trustworthiness when they know to be interacting with a “Disadvantaged”, probably because of within group preferences. Also, “Disadvantaged” senders increase their average level of trust when they interact with “Advantaged” receivers to earn higher profits from the transaction. Indeed, if reciprocity drives receivers’ decision to give something back to the sender, the positive variation of the endowment reduces the costs of giving so as to induce advantaged receivers to increase the level of transfers. “Disadvantaged” senders may anticipate this and they may increase their average transfer accordingly to earn higher profit from the transaction. A different behavioral pattern is observed from “Disadvantaged” receivers that increase their average level of trustworthiness when they know to interact with other disadvantaged players but decrease it when they are matched with “Advantaged” senders. Since only other-regarding preferences can motivate receivers choices to give something back, the information on the “Advantaged” state of the sender indicates a disparity in the monetary endowment inhibiting their level of reciprocity or altruism.

1.4 Structure of the thesis

This thesis is organized as follows. The first chapter describes the progressive development of a compounded experimental design aiming at eliciting individuals' preferences and beliefs of a non-standard kind that may be responsible for the emergence of cooperation in the trust game. The second chapter reports the results of an experiment identifying the systematic differences in trusting and trustworthy behavior amongst individuals' in different labor market states (i.e. employed, student or NEET) as well as the effect of information on others' labour market states over subjects' behavioral trust and trustworthiness. The third chapter delivers the results of an experiment that aims at identifying the effect of the random variation of subjects' endowment over their behavioral trust and trustworthiness in three different economic contexts as well as the way in which the information on the counterpart's state mediates subjects' trusting and trustworthy behavior. Finally, a brief conclusive chapter summarizes the results, emphasizes their relevant implications and proposes three main avenues for future research.

1.5 References

- Alesina, A., & Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature*, 53(4), 898-944.
- Alesina, A., & La Ferrara, E. (2002). Who trusts others?. *Journal of public economics*, 85(2), 207-234.
- Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3), 715-753.

- Anderson, L. R., Mellor, J. M., & Milyo, J. (2006). Induced heterogeneity in trust experiments. *Experimental Economics*, 9(3), 223-235.
- Andreoni, J., Harbaugh, W. T., & Vesterlund, L. (2008). Altruism in experiments. *The New Palgrave, Second Edition*, edited by Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan.
- Algan, Y., & Cahuc, P. (2010). Inherited trust and growth. *The American Economic Review* 100(5), 2060-2092.
- Arrow, K. J. (1974). *The limits of organization*. Norton, New York.
- Ashraf, N., Bohnet, I., & Piankov, N. (2006). Decomposing trust and trustworthiness. *Experimental Economics*, 9(3), 193-208.
- Baron, J. (2000). *Thinking and deciding*. Cambridge University Press.
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior*, 10(1), 122-142.
- Bohnet, I., Greig, F., Herrmann, B., & Zeckhauser, R. (2008). Betrayal aversion: Evidence from brazil, china, oman, switzerland, turkey, and the united states. *The American Economic Review*, 294-310.
- Bohnet, I., & Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior & Organization*, 55(4), 467-48
- Cox, J. C. (2004). How to identify trust and reciprocity. *Games and Economic Behavior*, 46(2), 260-281.
- Costa, D. L., & Kahn, M. E. (2003). Civic engagement and community heterogeneity: An economist's perspective. *Perspectives on politics*, 1(01), 103-111.
- Dearmon, J., & Grier, K. (2009). Trust and development. *Journal of Economic Behavior & Organization*, 71(2), 210-220.
- Eckel, C. C., & Wilson, R. K. (2004). Is trust a risky decision? *Journal of Economic Behavior & Organization*, 55(4), 447-465.

- Etang, A., Fielding, D., & Knowles, S. (2011). Does trust extend beyond the village? Experimental trust and social distance in Cameroon. *Experimental Economics*, 14(1), 15-35.
- Fehr, E., Fischbacher, U., Von Rosenbladt, B., Schupp, J., & Wagner, G. G. (2003). A nation-wide laboratory: examining trust and trustworthiness by integrating behavioral experiments into representative survey. IZA Discussion paper series, No. 715.
- Frank, R. H., & Sunstein, C. R. (2001). Cost-benefit analysis and relative position. *The University of Chicago Law Review*, 323-374.
- Gambetta, D. (2000). Can we trust trust?, in (D. Gambetta, ed.), *Trust: Making and Breaking Cooperative*, Basil Blackwell.
- Gauthier, D. P. (1986). *Morals by agreement*. Oxford university press.
- Guala, F. (2005). *The methodology of experimental economics*. Cambridge University Press.
- Guiso, L., Sapienza, P., & Zingales, L. (2009). Cultural biases in economic exchange?. *The Quarterly Journal of Economics*, 124(3), 1095-1131.
- Hardin, R. (2002). *Trust and trustworthiness*. Russell Sage Foundation.
- Heap, S. P. H., Tan, J. H., & Zizzo, D. J. (2013). Trust, inequality and the market. *Theory and Decision*, 74(3), 311-333.
- Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5), 865-889.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society*, 263-291.
- Karni, E., & Schmeidler, D. (2016). An expected utility theory for state-dependent preferences. *Theory and Decision*, 81(4), 467-478.

List, J., Sadoff, S., & Wagner, M. (2011). So you want to run an experiment, now what? An Introduction to Optimal Sample Arrangements. *Experimental Economics*, 14, 439-457.

Manapat, M. L., Nowak, M. A., & Rand, D. G. (2013). Information, irrationality, and the evolution of trust. *Journal of Economic Behavior & Organization*, 90, S57-S75.

Sapienza, P., Toldra-Simats, A., & Zingales, L. (2013). Understanding trust. *The Economic Journal*, 123(573), 1313-1332.

Smith, V. L. (1998). The two faces of Adam Smith. *Southern economic journal*, 2-19.

Thaler, R. H., & Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management science*, 36(6), 643-660.

Tsutsui, K., & Zizzo, D. J. (2014). Group status, minorities and trust. *Experimental Economics*, 17(2), 215-244.

CHAPTER 2: Experimental Investigations on Trust

2.1 Introduction

Experimental analyses generally test the empirical implications of Diego Gambetta's definition of trust and trustworthiness: "when we say that (...) someone is trustworthy, we implicitly mean that the probability that he will perform an action that is beneficial (...) is high enough for us to consider to engaging in some form of cooperation with him" (Gambetta 2000: 216). Trust is defined as the (possibly *rational*) belief of selfish individuals on others' reliability, while trustworthiness is intended to be a general disposition to reciprocate a trusting act with a non-detrimental choice to maximize individuals' self-interest in the long run. The definitions of trust and trustworthiness are complementary, because the existence of the general disposition to positive reciprocity justifies the subjective belief in others' reliability.

These definitions implicitly refer to anonymous or non anonymous interactions repeated in the long run where cooperation leads to a Pareto-superior outcome. Standard theory predicts trust to be sustainable as an equilibrium amongst selfish and rational players in anonymous interactions repeated in the long run, but not necessarily *ad infinitum*. In this case, the evidence of a low level of trustworthiness would induce the truster not to entrust others, so as to determine the emergence of sub-optimal outcomes. Anticipating this, the trustees have an incentive to provide evidence of trustworthiness so as to maximize their self-interest in the long-run. In the context of repeated interactions, knowledge of individuals' identity provides the trustees with a further reputational incentive to repay trust: When interaction are non-anonymous, the trusters have the chance to choose the most reliable counterpart, so

as to induce a competition amongst trustees that determines an increase of the average level of trustworthiness.

On this account, however, the emergence of trust and trustworthiness is related to the phenomenon of parochialism since individuals would tend to entrust or to be reciprocal with other individuals only within a restricted network. This kind of trust relationships acts as a constraint on the dynamics of impersonal trade in large markets, which consist of one-shot anonymous interactions whose efficient outcome strictly depends on individuals' trust. Hence, the main problem of standard theory is the explanation of the phenomenon of trust and trustworthiness in one-shot anonymous interactions, where rationality and self-interest prescribe individuals not to cooperate.

The motivations to entrust strangers in one-shot anonymous interactions consist of a beliefs and a preferences component. Rationality of beliefs would require trusters to understand the trustee's incentive to repay trust; this is clearly not the case of one-shot anonymous interactions. Hence, trusters may ground their expectations of trustworthiness on prior beliefs inherited through cultural transmission and update them adaptively by extending their past experience of trustworthy behavior in repeated interactions to the case of one-shot anonymous ones. Moreover, trusting and trustworthy behavior are compatible with hypotheses on other-regarding preferences that may derive from cultural transmission or biological evolution. Indeed, a trusting choice may stem from self-interest if the sender assesses the expected return to trust to be higher than the expected return to not trust. By the same token, trust may arise out of altruism or inequity aversion. Similarly, altruism, fairness and reciprocity may drive trustees' behavior. Both the belief and the preference component of trust, though not derivable from a normative theory of rationality, may coordinate individuals towards a Pareto-superior outcome in one-shot anonymous interactions.

The experimental analysis of trusting and trustworthy behavior is grounded in the so-called Trust Game (Berg et al. 1995), which represents the logic of sequential

transactions without enforceable contracts with anonymous participants, where behavioral trust and trustworthiness may arise. The game involves a sender (the truster) and a receiver (the trustee), both endowed with a monetary endowment. The sender chooses whether to give any percentage of it to the receiver. To incorporate positive return rate on trust, any amount sent is generally tripled and the receiver chooses whether to give back to the giver a percentage of the monetary amount actually received. Standard theory provides a prediction based on the assumptions of strategic rationality and selfishness: a rational sender will not send anything to the receiver because (s)he expects the latter not to give back anything to reap all the profits (if any) from the exchange.

This prediction is sharply in contrast with economic interactions in the field, where trusting and trustworthy behaviors are observed (e.g., Algan and Cahuc 2010). This gap leaves room for experimental investigations to identify the motivational determinants of trust and trustworthiness in one-shot anonymous interactions. In the last two decades, several contributions have provided convincing evidence falsifying the sub-perfect Nash-equilibrium solution of the trust game. Indeed, senders offer on average 50% of the endowment, while receivers send back 40% of the amount received (Johnson and Mislin 2011). The evidence of cooperative behavior in the trust game, however, is not enough to understand the determinants of behavioral trust and trustworthiness, which may involve different and non-exclusive motivations.

This chapter first clarifies how the beliefs and preferences component of trust and trustworthiness may have been selected by cultural or biological evolution, and then it illustrates the development of an experimental design suitable for reliable causal inferences from behavioral trust/trustworthiness to their motivational determinants. This protocol associates to the Trust Game different decision problems (either strategic or not), like the Dictator Game and lottery choices, to identify the idiosyncratic contributions of each of the multiple and non-exclusive motivations to the emergence of trusting and trustworthy behavior. Since the Nash equilibrium of the trust

game derives from the theory of strategic rationality and a set of assumptions about individuals' preferences and beliefs, this design allows understanding whether the phenomenon of cooperation in the trust game is due to the violation of rationality, of the assumptions of selfish preferences or to the fact that individuals' beliefs diverge from the standard kind.

The survey of the experimental literature clarifies that the phenomenon of cooperation in the trust game is mainly due to the violation of the auxiliary assumption of selfish preferences, as individuals systematically exhibit a trusting and trustworthy behavior based on (conditional or unconditional) other-regarding preferences. Furthermore, experimental trials provide evidence of a significant positive correlation of individuals' beliefs with their trusting behavior. However, the literature at issue does shed light on the mechanism of beliefs formation. In this regard, recent investigations (e.g. Fehr et al. 2003, Anderson et al. 2006) provide preliminary evidence that behavioral trust and trustworthiness vary across individuals' states (e.g. economic, social, family status). This evidence is compatible with two hypotheses. First, the explanatory relevance of different preferences (either selfish or other-regarding) changes in accordance with individuals' state. In this framework, individuals' state is a trigger of different kinds of preferences. Second, for a given type of preferences, individuals' state may influence the process of belief formation, so as to induce behavioral heterogeneity across states. The implications of these hypotheses are going to be explored in the next chapters.

The remainder of this chapter is organized as follows. Section 2.2. clarifies the distinction between trust and trustworthiness intended as rational behaviors, cultural traits and evolutionary traits. Section 2.3 illustrates the trust game. Section 2.4 surveys the main experimental analyses of the determinants of trust and trustworthiness. Section 2.5 describes the recent practice of combining survey and experimental data on trusting and trustworthy behavior. Section 2.6 highlights the issue of the

state contingency of trust and trustworthiness as a promising research line. Section 2.7 concludes.

2.2 What are trust and trustworthiness?

2.2.1 *Trust and trustworthiness as rational behaviors*

Trust is not a primitive but a concept reducible to a belief and a preference component (Fehr 2009). The belief component refers to the subjects' probability distribution about the trustee's behavior, while the preference component refers to both the individuals' interest to maximize their well-being and their risk attitude. Individuals' interest in well-being is not necessarily selfish and it may involve other-regarding preferences (i.e. preferences over one's own and other's material payoffs) such as altruism and a taste for fairness (Hardin 2002). Both the belief and the preference component can be expressed in a choice behavior - i.e. *trusting behavior* - which is a reliable proxy of individuals' level of trust.

Trust can be learnt as it requires the knowledge of trustee's reliability. The features of this learning strictly depend on the characterization of trustworthiness. A trustee may have a reputational incentive to repay trust in indefinitely repeated and non-anonymous interactions (Hardin 1992), or in finitely repeated games with a long enough time horizon (Benoit and Krishna 1987), where cooperation is optimal and can be sustained by rationality and self-interest. This account, however, does not explain why trustworthiness is observed in one-shot anonymous interactions. According to David Gauthier (1986), a trustee acts on the grounds of a disposition towards trustworthiness, which guides her behavior in a wide *class* of situations and not *case-by-case*. In this view, a disposition is rational if and only if it grants to the individual the highest general utility in the relevant class of situations. Hence, the disposition-based account of trustworthiness predicts a stability in behavior, entailing the possi-

bility to repay trust even in one-shot anonymous interactions. As in the case of trust, trustee's behavior is a reliable proxy of her disposition towards trustworthiness.

The disposition-based account of trustworthiness entails that the knowledge that the truster stores in repeated (either anonymous or non-anonymous) interactions is relevant even in one-shot anonymous interactions. By the same token, the disposition toward trustworthiness is valuable to the trustee only if the truster has at least an imperfect knowledge of it (Bruni and Sugden 2000). In one-shot anonymous interactions, truster's expectations are just a generalization of past experiences, on the basis of which she can estimate the likelihood of the trustworthiness of a *typical person* in the trustee's position. This estimation is the result of the truster's inference from its context of reference. Thus, the identification of the level of trustworthiness in a given social context is the precondition of a trusting choice toward an absolute stranger. The material consequences of the one-shot anonymous interaction are stored as new information on the average level of trustworthiness in a given context.

The updating mechanism of beliefs is an open issue. According to Hardin (2002), individuals update their prior probabilities on trustees' reliability as naive Bayesians at any iteration of trust relevant interactions so as to learn their optimal level of trust on the grounds of their personal history. However, trusting behavior is observed even when there is no evidence (or counter-evidence) on the trustee's reliability. This suggests that trusters' expectations about others' trustworthiness may systematically violate the requirement of Bayesian updating. The next section tries to further clarify this issue.

2.2.2 *Trust and trustworthiness as cultural traits*

Trust and trustworthiness are generally intended to be cultural traits shared within a given society. Being the result of a systematic interaction between preferences and expectations about others' trustworthiness, trust is consistent with a definition of

culture in terms beliefs and values. According to Bisin and Verdier (2010), culture consists of a system of beliefs and preferences about the consequences of individuals' actions that are transmitted through interaction within and across generations.

Cultural transmission shapes basic primitives of trust such as risk attitudes and altruism, together with individuals' prior expectations about others' trustworthiness. Specifically, children can learn their prior beliefs about others' trustworthiness as well as the value of reciprocity from the ongoing relationship with their parents (*vertical socialization*). Indeed, parents create a sense of trust in the child by providing an unambiguous example of personal trustworthiness. Moreover, this example induces into the child a *disposition* for reciprocal trust that guides her behavior in a wide class of interactions. Hence, the formation of beliefs about parents' trustworthiness (i.e. trust) is not independent from the formation of the personal disposition to trustworthiness. Put differently, the solution to the developmental problem of whether to trust parents or not should come together with the solution to the developmental problem of entrusting oneself. Because of that, the choice of entrusting others is at the roots of the development of the ego.

Children can integrate the cultural traits inherited from parents through imitation and social learning in a given context (*oblique socialization*). Economic models of cultural transmission are based on the realistic assumption of parents preferring their children to adopt their own cultural traits. According to Bisin and Verdier (2010), if vertical and oblique socialization are cultural substitutes or complements, the model generates respectively a stationary distribution of heterogeneous or homogeneous cultural traits. Specifically, in the case of cultural substitutability, families socialize their children more intensively when their cultural traits are less frequent in the population. Hence, cultural substitutability preserves the heterogeneity of cultural traits. Conversely, cultural complementarity entails that families socialize their children more intensively when their cultural trait is dominant in a given social context. This leads in the long run to an homogeneous distribution of cultural traits. This ac-

count guarantees inter-generational and between-subjects heterogeneity in the capacity to trust or to be trustworthy. Indeed, individuals may differ in their attitudes toward risk and altruism as well as in their subjective beliefs about others' trustworthiness.

However, the persistence of different attitudes towards trust and trustworthiness requires an additional assumption about the learning process. The general relevance of the information about trust and trustworthiness acquired from parents entails an extension of the knowledge stored in repeated non-anonymous relationships to other kinds of interactions involving trust. Recent theoretical and empirical contributions in cognitive sciences demonstrate that when children learn to trust, they tend to treat the piece of available information on others' trustworthiness as *representative* of the general level of trustworthiness in a given social category (e.g. group membership). For example, a child who gets used to exemplary trustworthy parents, and therefore develops a strong disposition toward trustworthiness, is more likely to entrust a stranger of the same group in an anonymous one-shot interactions (Landrum et al. 2015). This process can be thought of as an heuristic that simplifies the cognitive task of entrusting strangers and dramatically reduces the uncertainty inherent to one-shot anonymous interactions involving trust.

This account entails that individuals' beliefs about the counterpart's trustworthiness in one-shot anonymous interactions do not form on the grounds of the knowledge of the trustee's incentives to repay trust. In this sense, trusters' beliefs are not rational. Indeed, if individuals (from early childhood) treat a piece of information on others' trustworthiness as representative of an entire social category, their beliefs can be quite stable and difficult to update as new information accrue from trust relevant interactions. Recent evidence shows that children continue to entrust subjects who proved several times to give them wrong information (Vanderbild et al. 2014). More generally, a marked difference across individuals' beliefs is observed even if they are

faced with the very same objective level of trustworthiness in a given social context (Bigoni et al. 2017).

Finally, given that the formation of beliefs about others' reliability is not independent from the development of a personal disposition to trustworthiness, individuals are likely to ground in introspection their expectations of others' behavior in one-shot anonymous interactions. On this account, if the distribution of types in the population (e.g. opportunistic, trustworthy and altruistic) is not observable - as it is the case in one-shot anonymous interactions - beliefs about others' trustworthiness are type-dependent: i.e. Individuals expect others to behave as they would if they were in others' position (Adriani and Sonderegger 2015).

2.2.3 Trust and Trustworthiness as evolutionary traits

Generally, economic models (e.g. Fehr and Schmidt 1999; Charness and Rabin 2002; Dufwenberg and Kirchsteiger 2004) provide only proximate explanations of cooperative behavior grounded in the assumption of the existence of other-regarding preferences (e.g. altruism, inequity aversion and reciprocity), but they close in a black box the ultimate explanation of the evolutionary mechanism favoring their selection.

A recent evolutionary model of trust and trustworthiness (Manapat et al. 2012) aims at delivering an ultimate explanation of the experimental evidence of cooperation in one-shot anonymous trust games. The main idea of this model is that individuals' other-regarding preferences proved to be adaptive in broad contexts and in long periods of time; hence, the cooperative behavior observed in one-shot anonymous interactions is nothing but the byproduct of the evolution of other-regarding preferences emerging from the repetition of anonymous or non-anonymous trust games in the long run.

The model considers a population of trusters and trustees repeatedly playing the trust game. On the one hand, the trusters make a transfer to the trustee with proba-

bility p in an anonymous interaction, while the trustees choose whether to return a fraction of the received amount. The trustors face two possible scenarios: (1) With probability $1 - q$ they do not get information on the trustee's level of trustworthiness (anonymous interaction); (2) With probability q , they get information on the fraction the trustee is willing to return before the interaction occurs. In this setting, trustees do not know the scenario the truster is facing, so they have a fixed strategy across all interactions. In this model, both players cumulate their gains through the repetition of the interaction. Moreover, trustors can choose the partner with whom engaging in a transaction conditional on their knowledge. This feature induces a competition amongst trustees.

The main result of this model is that even a small probability of getting information on the counterpart determines a distribution of trusting and trustworthy individuals. Furthermore, the possibility of partner choice induces a substantial increase in the level of trustworthiness in the population.

This model is based on the assumption that in the long run period of evolution, the distinction between choices and preferences blurs. Hence, strategic choices repeated in the long run shape more fundamental traits such as preferences, which are generally taken as fixed in economic models. On this account, one-shot anonymous interactions can trigger individuals' other-regarding preferences, which proved to be adaptive in the long run. Moreover, the authors claim that their model can be interpreted either as genetic evolution or as a process of social learning. The former interpretation entails a definition of trust as a biological primitive. However, it is not clear whether a genetic account of trust amounts to be a biological reduction of trust as a cultural trait or rational behavior. The neuroeconomic evidence on the phenomenon of cooperation in personal interactions is far from conclusive: i.e. experiments on the neural correlates of cooperative behavior produce non-robust results (van Rooij and Van Orden 2011). Thus, it is reasonable to conjecture that genetic evolution determined the development of preferences and feelings (e.g. altruism and guilt), which

are automatically triggered in trust relevant interactions so as to work as enforcement devices of the individuals' *deliberate* choice of implementing a Pareto-superior outcome. This means that the behavioral relevance of other-regarding preferences such as altruism in relationships involving trust allows to reach a mutually beneficial outcome independently of any *belief* about the counterpart's trustworthiness. The empirical implication is that a lower level of cooperation would be observed in one-shot anonymous trust games if individuals' behavior was uniquely based on selfish preferences and beliefs about others' trustworthiness.

2.2.4 Young people's trust and its consequences for the labor market

As learnt capacities, trust and trustworthiness are naturally sensitive to critical periods of personality development. In early childhood, parental care and the family environment in general are the first examples of trustworthy behaviors that provide an imprinting over the individual capacity to recognize others' reliability and to repay others' trust. In this light, it comes as no surprise that children in conditions of neglect or that suffered parents' abuse in early age systematically exhibit lower levels of trust and trustworthiness (Alesina and Giuliano 2015). The evident affective implications of the parents-children relationship entails the transmission of a more fundamental trait of personality (i.e. preferences) and not only of a set of beliefs.

Beside the above mentioned affective investments, parents are called to invest in their children's human capital. Such investment is likely to depend on both parents' and children's level of generalized trust, which positively corresponds to the levels of trust in cultural institutions (Putnam 1993): i.e. higher trusting parents are likely to invest more in education because of their higher level of trust in cultural institutions. Investments in human capital exploit the cooperative nature of human culture, where adults teach altruistically young people and the latter conform to adults' in order to fit cooperatively in the relevant cultural group (Tomasello 2014). Hence, in-

investments in human capital support an ongoing dynamics of cooperation, in which trust and trustworthiness play a crucial role. Moreover, culture allows young people to construct self-knowledge in relation to the relevant cultural group and by assessing one's own difference with respect to other cultural groups. This entails that culture fosters the higher level capacity of imagining oneself in the situation of different persons. This capacity is crucial to learn how to identify others' reliability as well as to repay others' trust, particularly in one-shot anonymous interactions. More generally, investments in human capital and the consequent development of the capacity of putting oneself in others' shoes are likely to mitigate the within-group bias deriving from the cognitive heuristics described in 2.2.2.

Human capital investments, as well as their direct effect on trust and trustworthiness, have long lasting consequences for individuals' success in the labor market. As widely recognized, well-educated people find jobs faster and get fired less frequently compared to less educated people (Cairò and Cajner 2016). Indeed, if higher investments in human capital foster trust, more trusting individuals are more likely to engage in mutually beneficial interactions, and for this reason they have a higher success rate in the job search with respect to less educated people. Moreover, if well-educated people are more trusting, than they are more likely to be cooperative in the work place, so as to increase the firm's benefits and reduce the probability of being fired. This success in the job market is likely to boost further people's willingness to entrust others or to repay their trust: i.e. if trust affects individuals' labor market state, the latter, in turn, influences individuals' level of trust and trustworthiness. A symmetric argument holds for less educated people. Hence, employed young people should exhibit systematically higher levels of trust and trustworthiness compared to young people that are unemployed or in precarious forms of employment.

2.3 The trust game

To identify the determinants of trusting and trustworthy behavior in one-shot anonymous interactions, the experimental analysis make us of the so-called trust game (Berg et al. 1995). Consider a dyadic sequential transaction without enforceable contracts between two individuals. The interaction consists of the following steps.

- Nature assigns to two players the role of sender (truster) and receiver (trustee). The two players do not know each other.
- Both players are endowed with the same monetary amount and the sender chooses whether to give a percentage of it to the receiver. The amount sent to the receiver – *if any* – is tripled.
- The receiver chooses whether to give back to the sender a percentage of the received amount.

This is a sequential game of perfect and complete information whose sub-game perfect Nash-equilibrium can be identified by means of three assumptions.

Assumption 1 (*Rationality*): players exhibit coherent choices based on consistent preferences and beliefs.

Assumption 2 (*Common knowledge of rationality*): Players are rational, players know that other players are rational, players know that the other players know that they are rational, and so on *ad infinitum*.

Assumption 3 (*Self-interest*): Both players have selfish preferences for they prefer more money to less without caring about other's payoffs.

If these assumptions are satisfied, the sender believes that at the second stage the receiver will not give back anything, and as a consequence at the first stage she will keep all the endowment for herself. Therefore, the sub-game perfect Nash-equilibrium prediction excludes trusting and trustworthy behavior. The interaction could lead to a Pareto-improvement if and only if for any tripled amount sent, the second

player gives back more than the original amount sent by the sender (Berg et al. 1995). However, this result is not sustainable as an equilibrium. Yet, several experimental investigations of the trust game provide robust evidence that is inconsistent with the sub-game perfect Nash-equilibrium prediction. Indeed, in experimental trust games senders on average offer 50% of the endowment, while the receiver gives back to the sender approximately 40% of what she got (Johnson and Mislin 2011).

Observed behavior has been interpreted in terms of trust and trustworthiness. On the one hand, the choice of the sender to give a positive amount at the first stage entails that she entrusts (i.e. expects) the receiver not to take all the money. On the other hand, the receiver can interpret the sender's choice as the willingness to entrust her to reach a Pareto-superior outcome at the second stage. This interpretation is likely to trigger the receiver's disposition to reciprocity, inducing her to implement the trustworthy choice (Smith 1998).

Nonetheless, the empirical evidence at stake raises the issue of what has been *exactly* falsified of the complex deductive machinery that leads to the sub-game perfect Nash-equilibrium. Given the simplicity of the trust game, it is doubtful that the available evidence falsifies the assumption that subjects are coherently acting on the grounds of their preferences, and on their beliefs in the case of senders. Therefore, individuals in the trust game probably behave in accordance with a minimal requirement of rationality. Conversely, it is more intuitive to think that the trust game falls in that class of games (e.g. Ultimatum Game and Dictator Game) where interpersonal comparisons might induce individuals to reveal a kind of preference that violates the assumption of selfishness. This violation entails that the players of the trust game might be of *different types*, e.g. selfish, altruistic and reciprocal. This heterogeneity introduces *uncertainty* about the payoffs function of the other player and legitimates different beliefs about her behavior. Therefore, the phenomenon of cooperation in the trust game entails a problem of empirical identification of the idiosyncratic contribution of non-standard preferences and beliefs as motivating factors

of trusting and trustworthy behavior. The following section illustrates the main experiments tackling this issue.

2.4 Determinants of trust and trustworthiness

This section describes an experimental protocol that associates to the trust game other decision problems (either strategic or not), to distinguish between competing hypotheses on individuals' behaviour. The abstract form of this problem can be described as follows. In the experimental design α two hypothetical causes x and y might be at work, but the behavioural data the design generates do not distinguish between them. A possible way to address this problem is to associate to the original design α a design β , which is identical to α in all respects except for the fact that one of the two hypothetical causes, let us say x , is "switched off". The evidence produced in this new design is primarily a test of the hypothesis y in the design β . However, given the similarity between α and β , the evidence supporting x in β is a legitimate basis to make inferences about the relevance of y in α , where x is "turned on". On these grounds, the variation of the phenomenon observed in design α with respect to β is imputed to x .

2.4.1 *Trust and trustworthiness in the lab*

Berg et al. (1995) provide one of the first experimental tests of the hypothesis of cooperative behavior in the trust game. The authors assume that receivers are likely to interpret any positive amount sent as an attempt of senders to use trust to reach a mutually beneficial outcome. This interpretation induces receivers to engage in reciprocal behavior: i.e. they repay trust by sending back a positive amount of money, which is increasing in the level of transfers from the sender. This hypothesis legitimates senders' beliefs on receivers' reliability; specifically, senders expect the return

from the transaction to be increasing in their level of transfers. In this sense, trust is a coordination device on a Pareto-superior outcome. Berg et al. violate the assumption of selfish preferences only in characterizing receiver's behavior. Indeed, if the sender assesses the expected utility of trusting to be higher than the expected utility of not trusting, then she will implement a cooperative behavior. This account is, however, problematic because according to the hypothesis of reciprocity a utility maximizing sender would send the whole amount on the grounds of the expectation of the highest return from the transaction. This is, however, inconsistent with the evidence of the average level of transfers in one-shot anonymous trust games (Johnson and Mislin 2011).

According to Berg et al. social history (i.e. public information on the level of trust and trustworthiness in the experimental sample) provides a support to both the sender's beliefs on the trustee's reliability and the receiver's interpretation of the trusting choice as an attempt to reach a Pareto-superior outcome. This hypothesis is based on the assumption that individuals' willingness to trust or to repay trust strictly depends on their socialization: i.e. the construction of individual's social identity through the identification with a group. This entails that the social history treatment boosts trust and trustworthiness by triggering individuals' identification with a group. To test these hypotheses, Berg et al. implement an experimental design where two samples of students of the University of Minnesota, all endowed with the same monetary amount, play a one-shot trust game in two different informational conditions. In the control group, subjects play a trust game with no information. In the treatment, participants play the very same game, but they receive historical information on the level of senders and receivers' transfers in the control group. The information at issue is supposed to boost the average rate of cooperation in the treatment group.

The results of the experiment confirm the interpretation of the cooperative behavior in the trust game in terms of reciprocity. In the control group, a significant share of

senders transfer to the receivers approximately 50% of their endowment. Receivers exhibit a reciprocal behavior whenever they get an amount between 50-100% of the sender's original endowment. Furthermore, for particularly high levels of sender's transfers, receivers send back a higher amount of money, so as to determine a net profit. The provision of social history information slightly increases the average level of senders' trusting behavior, while it determines a *significant* increase in the average transfer from receivers.

Berg et al.'s experiment, though providing empirical support to the hypothesis of cooperative behavior in the trust game, does not identify the specific motivating factors leading subjects to transfer a positive amount of money. Senders' choices are consistent with both the hypothesis of satisfaction of selfish preferences given their beliefs about receiver's trustworthiness and with the hypothesis of unconditional other-regarding preferences (e.g. altruism and inequity aversion) where beliefs play no role. Analogously, receivers' choices are consistent with both the hypotheses of reciprocity and unconditional other-regarding preferences. Therefore, the single trust game is of no use to address this problem.

2.4.2 *Trust and trustworthiness are not pure altruism*

Cox (2004)'s experiment aims at distinguishing pure altruism from trust and trustworthiness. There is a general agreement in defining altruism as the willingness to act in the interests of other persons, without any further motivation (Andreoni *et al.* 2008). To identify the different behavioural consequences between trust/trustworthiness and altruism, Cox implements a *between-subjects* experimental design consisting of two dictator games and a trust game. The dictator game is similar to the trust game but the receiver does not take any decision and the game ends with the sender's choice on the share of the endowment to be given. The dictator game is an individual (non-interactive) decision problem where strategic motivations to trust

(i.e. the *expectation* of a return based on the reliability of the receiver) as well as the propensity to reciprocate are not at work. Hence, any positive amount sent in the dictator game is due to unconditional other-regarding preferences. Cox compares the average behaviours of the two different samples elicited in the trust and dictator games to identify the effect of trust and reciprocity as distinguished from altruism.

Coherently with Berg et al. (1995), Cox assumes trust to stem from the giver's aspiration to realize a higher pay off by running the risk of receiver's defection. This is consistent with the definition of trust as the expectation to realize gains from the transaction with a trustworthy receiver. Since this possibility is neutralized in the dictator game, the author considers trust to be an *additional* motivation with respect to altruism. This means that trust should induce the sender in the trust game to give a higher amount of money than the amount sent in the dictator game on the grounds of altruistic preferences. Thus, according to Cox altruistic preferences determine the default transfer, while trust (i.e. the expectation of others' trustworthiness) induces only positive variations of this baseline. In this view, trust is distinguishable from unconditional other-regarding preferences by simply computing the difference between the average transfers from givers and dictators.

In a similar fashion, to distinguish between the hypotheses of trustworthiness and altruism, Cox compares the average behavior of receivers in the trust game with the average receivers' behavior in a modified dictator game. To neutralize any reason for reciprocal behavior, the allocation of money between senders and receivers should be random, so as to elicit receivers' pure altruism. To pursue this aim, Cox designs a modified dictator game where senders and receivers are randomly associated to a sender of the first trust game. Senders get the amount not sent by the first mover in the trust game, while receivers get the amount given by the sender in the trust game. Hence, receivers play the role of dictators and have the chance to choose the amount of money to give to the sender. Since the allocation of the endowment is *exogenously* determined by a random mechanism, this design switches off any feeling of obliga-

tion towards the senders that might motivate receivers' reciprocal behaviour. In this context, receivers' transfers are a measure of altruism: i.e. the default transfer from receivers in absence of any generous action from the sender. According to Cox, a trustworthy behavior in the standard trust game entails an *additional* cost with respect to the altruistic behavior elicited in the modified dictator game, "in the sense that the amount returned is larger than the amount that would maximize the second mover's utility in absence of the generous action by the first mover" (Cox 2004: 268). Hence, reciprocity is computable as the positive difference between the average transfer from receivers in the trust game and in the modified dictator game.

Cox's experiment delivers empirical support to these priors. Indeed, the average investments of senders are higher in the trust game than in the dictator game, providing evidence of trust as distinguished from pure altruism. Similarly, the positive difference between the average receivers' transfers in the trust game and the average receivers' transfers in the modified dictator game is an evidence of positive reciprocity as distinguished from pure altruism. This evidence provides the basis to senders' expectation of receivers' actions that implement a Pareto-improvement with respect to the sub-optimal outcome of the Nash-equilibrium solution.

2.4.3 Distinguishing trust and trustworthiness from unconditional other-regarding preferences

The results accruing from Cox's experiment provide only limited evidence for the explanatory hypotheses on trust and trustworthiness. Indeed, the author defines trust as the upshot of a calculation that takes into account the expected utility of the states of the world consequent to the choice of trusting or not trusting. However, Cox's experiment does not elicit senders' beliefs about receivers' level of transfers, which are crucial to distinguish utility maximizing behaviour from behavioral regularities grounded in unconditional other-regarding preferences. Furthermore, the

characteristics of receivers' return function (i.e. how much the receiver gives back for any amount received) are important to distinguish positive reciprocity from different kinds of other-regarding preferences that might regard the distribution of wealth (e.g. inequity aversion). Cox's design elicits receiver's *single* choice contingent on sender's transfer; because of that, it does not provide any evidence of such return function. Finally, the most controversial hypothesis by Cox is that both trust and trustworthiness are *positive* differences between the net transfers in the trust and dictator game. This account neglects the possibility that receivers' active role in the trust game might trigger behavioural attitudes (e.g. betrayal aversion) that make hard to predict the sign of the variation of players' transfers across the dictator and trust game.

To fine-grain the evidence on the motivations behind the behavioural regularity of positive net transfers between givers and receivers, Ashraf *et al.* (2006) propose an experimental design based on the comparison of subjects' choices in dictator and trust games. More precisely, Ashraf et al. implement a standard Dictator Game and a Tripled Dictator Game, with the difference being that in the second game any amount given by the dictator is tripled. With respect to Cox's, Ashraf et al.'s design exhibits three main features of innovation. First, the authors propose a *within-subjects design*, where the *very same* sample of individuals play the dictator and trust game. This design controls for individuals' heterogeneity in their (social) preferences and it allows identifying the correlation between subjects' choices in the dictator and trust game without any assumption of additionality of trust and trustworthiness with respect to altruism. Second, senders' are asked to state (without reward for accuracy) a point estimate on the level of transfers from the sender. Third, the authors implement the so-called strategy method, where the sender chooses the amount to be sent in the standard way, while the receiver has to choose the amount to be sent back conditional on the potential amount sent by the sender, without knowing the actual choice of the first player.

Ashraf et al. (2006)'s results on trusting behavior are consistent with Cox's experimental evidence. Both expectations about receivers' trustworthiness and unconditional other-regarding preferences elicited in the dictator games are economically and statistically significant coefficients. Unconditional other-regarding preferences are here treated as a control variable and not as the baseline with respect to which trust is estimated as a positive variation. Therefore, in Ashraf et al.'s experiment the significance of the coefficient of unconditional other-regarding preferences does not necessarily stand as a support of Cox's assumption of additionality of trust.

The results on receivers' behaviour are consistent with the empirical implications of the hypothesis of reciprocity, as the authors find a positive relationship between senders' level of transfers and the amount sent back by the receiver. Nevertheless, this positive relationship might be due to individuals' preferences for the distribution of money, intended as the trustee's interest in the *final* (and possibly *equitable*) allocation of the monetary amount, which does not entail the trustee's willingness to repay trust (or distrust) by being kind (or nasty) in return. To distinguish between the two hypotheses, the authors run two different regressions: the first one aims at controlling for the significance of unconditional other-regarding preferences (elicited in the dictator game) together with positive reciprocity (elicited in the trust game), while the second regression replaces the coefficient of unconditional other-regarding preferences with a model of fixed preferences for the distribution of money. The result of the first regression shows that both positive reciprocity and unconditional kindness are statistically and economically significant in explaining trustworthy behavior. However, in the second regression the effect of positive reciprocity disappears. The authors conclude that the positive slope of receivers' return function is mainly due to their preferences for the distribution of the monetary endowment and that positive reciprocity plays a relatively minor role.

2.4.4 Trusting behavior and individuals' attitude towards risk

A further source of individuals' heterogeneity in experimental investigations of trust is their attitude towards risk. It is intuitive to think of the trust game as an investment game, where the sender's attitude towards the probability of receiver's trustworthiness is the crucial determinant of trusting behavior.

To control for the influence of risk attitudes over trusting behavior, subjects are involved in individual decision problems where they make several choices among a series of lotteries with increasing levels of risk. The aim of this design is to predict trusting behavior from individuals' propensity towards risk, under the hypothesis that the more risk averse a subject is, the lower her level of trust is. This approach has been proposed in several variants (e.g. Eckel and Wilson 2004, Ashraf et al. 2006, Houser et al. 2010), but it delivers the negative result of non robust correlation between the individuals' risk attitudes and trusting behavior.

The experimental analysis of trust involves inferences from the behavioral task of *control* (e.g. risky decision problems) to the *target* behavioral task (e.g. the trust game). The validity of these inferences crucially depends on the similarity between the two tasks. However, epistemic and ontological reasons entail that the data accruing from the individual decision problem can hardly be a measure of subjects' trusting behavior. First, the trust game is a choice problem in conditions of uncertainty with *imperfect* knowledge of the likelihood of the receiver's decision; conversely, in the standard model of risky decision problems individuals have *perfect* information on the likelihood of the realization of any given outcome independently from the decision of other persons. This epistemic difference entails that individuals' choices in risky decision problems do not fully capture the formation of individuals' beliefs in conditions of uncertainty. Second, strategic interactions involve *persons*, while risky decision problems involve a random *mechanism* generating the outcome. This onto-

logical difference entails that in the trust game the interaction with a human being might trigger behavioural attitudes that are off work in risky decision problems.

To address these problems, Bohnet and Zeckhauser (2004) and Bohnet et al. (2008) propose the alternative hypothesis of betrayal aversion defined as the disutility of experiencing or anticipating receiver's cheating behavior. Betrayal aversion is an *additional* non financial costs that further inhibits trust with respect to standard risk aversion. According to this hypothesis, when individuals face the very same payoffs and probabilities in a risky decision problem and in a strategic interaction, they are more willing to take risk when the outcome is determined by a random mechanism than to trust another person that might cheat on them.

To test this hypothesis, Bohnet and Zeckhauser (2004) propose an experimental design involving a trust game and a risky decision problem with the very same *binary* payoff structure. In the trust game, the sender has to choose between the certain amounts of her endowment of 10 tokens and a lottery, whose outcome is determined by the receiver. If the second player chooses to be trustworthy, the game ends with the payoffs pair (15, 15), while if the receiver chooses to cheat, the game ends with the payoff pair (8, 22). The authors elicit senders' minimum acceptance probability (MAP), which is the minimum percentage of receiver's trustworthy behaviours in the experimental group *required* by the sender as a necessary condition to implement the trustful choice. The actual percentage of trustworthy behaviour is determined by eliciting receivers' choices through the strategy method. The information about the percentage of trustworthy receivers is communicated *after* the senders have revealed their MAP. If the actual percentage of trustworthy behavior is equal or higher than MAP, then the senders' trusting choices are automatically implemented; the opposite holds, if the actual percentage of trustworthy behavior is lower than MAP. Moreover, Bohnet and Zeckhauser elicit individuals' propensity towards risk in risky decision problems involving a sure payoff of 10 or a lottery $L = \{15 \text{ with probability } p \text{ and } 8 \text{ with } (1-p)\}$. As in the trust game, the authors elicited the individuals' MAP, which is

the minimum probability of the favorable outcome (15) that would induce the individual to choose the lottery over the certain outcome. In this risky decision problem, $2/7$ represents the MAP of risk neutral individuals. This allows to easily identify risk averse (i.e. $\text{MAP} > 2/7$) and risk loving subjects (i.e. $\text{MAP} < 2/7$). In this experimental task, the actual value of p is determined by a random mechanism at the outset. Given the characterization of betrayal aversion as an additional inhibiting factor of trusting behavior, if the MAP elicited in the trust game is higher than the MAP elicited in risky decision problems, then betrayal aversion is a relevant motivating factor. The results of Bohnet and Zeckhauser's (2004) experiment supports this hypothesis, since they observe a higher MAP in the trust game than in the risky decision problems.

This experiment does not control for social preferences as possible determinants of individuals' behavior in the trust game. In a further experiment, Bohnet et al. (2008) compare the MAP levels in a binary trust game, with those elicited in a modified dictator game, where the receiver still plays a passive role, but a random mechanism determines the amount to be sent back to the giver. This is a way to transform the trust game in a risky decision problem without neutralizing the social factors inherent to personal interactions. In this design, individuals' social preferences could still be active because the receiver is a human being, but this should be not betrayal aversion since the receiver plays a passive role. Thus, Bohnet et al. considered the positive difference between MAP in the trust game and the MAP in the modified dictator game as a more reliable evidence of betrayal aversion. The authors found strong evidence in support of the betrayal aversion hypothesis, so to confirm the behavioural relevance of the distinction between strategic interactions in conditions of uncertainty (i.e. trust game) and individual decision problems under risk (i.e. lottery choice).

2.5 Correlation between survey and experimental data on trust and trustworthiness

As emphasized in 2.2.2, individuals learn to trust others in the real world by repeatedly engaging in social interactions (either anonymous or non-anonymous) where the counterpart has the discretionary power to lower (or increase) their utility. The crucial implication is that individuals' past experiences and their socio-economic characteristics determine their heterogeneity in trusting behavior in anonymous one-shot interactions. A similar argument holds for trustees as behavioral trustworthiness is likely to differ according to the individual experience of trusting behavior from other people. However, the experiments described in the previous section do not control for the way individuals' behavior in the lab relates to their socio-economic characteristics and general attitude towards trust/trustworthiness.

To fill this gap, several empirical analyses combine experimental investigations on trust with survey methods providing detailed information on the characteristics of individuals' beliefs and preferences, not inferable from behavioral data. This approach uses the standard experimental analysis of trust to measure non-experimental phenomena elicited through surveys. Indeed, the experimental evidence validates the behavioural content of subjects' answers to survey questions. Therefore, the combination of survey and experimental data aims at increasing the internal validity of the empirical test of behavioural hypotheses on trust and trustworthiness.

The prototypical experiment of this kind is articulated in two phases. First, subjects are required to answer to a series of questions aimed at eliciting their socio-economics characteristics, their general attitude towards trust and their experience of past trusting behaviours. The questions on trust are taken from the questionnaire of the World Value Survey (WVS). An example of attitudinal question on trust is the following: "Generally speaking, would you say that most people can be trusted or that you cannot be too careful in dealing with people?" This question is aimed to reconstruct the individual's beliefs on people's reliability and it is generally associated

to questions on the experience of trusting behaviors of the following kind: "Have you ever spontaneously benefited from a person you did not know before?" The second part of the experiment provides behavioural evidence of trust and trustworthiness; therefore, it consists of a trust game and a series of other experimental tasks, which might vary in accordance with the specific control variables of interest (e.g. risk aversion and altruism). Generally, the survey part is conducted some time before the behavioural part to mitigate the problem of experimenter demand effect (i.e. the change in behaviour due to some cue suggesting the "appropriate" behaviour in the experiment).

Glaeser et al. (2000) provide one of the first contributions based on the integration of survey and experimental data on trust and trustworthiness. In the first phase of the experiment, subjects are required to answer attitudinal and behavioural questions on trust, while in the second phase they play a standard trust game. The authors identify a significant correlation between the answers on past experiences of trusting behaviours and behavioural trust elicited in the lab. However, they find a significant correlation of subjects' answers to WVS attitudinal question with behavioral trustworthiness but not with trust.

Fehr et al. (2003) refine the analysis of the correlation between survey and experimental data with an experimental design that exhibits two relevant features of innovation with respect to Glaeser et al. (2000). First, they run a nationwide experiment with a representative sample of German people to avoid the problem of homogeneous sampling. Second, the authors elicit the beliefs of subjects playing the role of senders in the trust game. The main result is that subjects' answers to attitudinal questions are predictive of behavioural trust *only if* the regression model do not control for senders' beliefs about the level of transfers from receivers. Thus, individuals' attitudes towards trust become operative via individuals' beliefs about receivers' trustworthiness. In contrast with Glaeser et al., the authors do not find any signifi-

cant correlation between the answers to attitudinal questions on trust and behavioural trustworthiness.

The comparison between Glaeser et al. (2000) and Fehr et al. (2003) shows that the evidence on the correlation between attitudinal questions on trust/trustworthiness and the trusting/trustworthy behaviour elicited in the lab exhibits a clear problem of robustness. To tackle this issue, Sapienza et al. (2013) propose an experimental design, where participants play *both* roles of sender and receiver in the trust game. The general result is that the WVS attitudinal question on trust stands in a significant correlation with senders' beliefs about receivers' level of trustworthiness. The WVS question is significantly correlated to senders' behavior, but the significance of this correlation dramatically decreases when regression analysis controls for senders' expectations about receivers' trustworthiness. This result is a hint for a correlation between WVS attitudinal question and subjects' beliefs, which in fact results to be significant. Thus, Sapienza et al. conclude that the WVS attitudinal question is a reliable measure of the belief component of trust.

Moreover, the authors find a significant correlation between individuals' beliefs when they play in the role of senders and the amount of money sent back when they play as receivers. Probably, in anonymous one-shot interactions senders integrate their prior probability on the reliability of the other player through a counterfactual reasoning that allows them to view the decision problem from the perspective of the receiver. Intuitively, in determining the level of transfer, senders ask themselves the amount of money they would send back, were they playing the role of receiver. This is consistent with the account of trust as a cultural trait, which entails that children's learning from parents form their expectations on others' trustworthiness together with the personal disposition to trustworthy behavior. Thus, in the case of the trust game, subjects' beliefs as senders are relevant motivating factors determining their choice as receivers. Moreover, the correlation of the answers to the WVS attitudinal question to senders' beliefs together with the correlation of senders' beliefs to their

behavior as receivers, entail a correlation between subjects' answers to the WVS attitudinal question and trustworthy behavior, which indeed results to be significant. Overall, this evidence suggests the relevance of introspection as a candidate mechanism of beliefs formation in decision problems under uncertainty.

2.6 State-contingent trust and trustworthiness

The research program based on the comparison of survey and experimental data raises the interesting issue of the dependence of behavioural trust and trustworthiness from the state of the individual. Individuals' state is here used as a general concept that subsumes several empirical instances such as social, labour and family status. It is therefore natural to conjecture that individuals' state might affect the motivating factors (e.g. self-interest, altruism, risk attitude, reciprocity and beliefs) of trusting and trustworthy behavior.

The state of the art of this possible research line is, however, at a very preliminary stage. Indeed, the relevant experiments identify exclusively a correlation between subjects' state – elicited through surveys - and their trusting/trustworthy behaviour. The interpretation of these results in causal terms would be fallacious, since the variables inherent to the subjects' state are not exogenous.

To illustrate the point, consider Fehr et al. (2003)'s experiment that identifies a negative effect of unemployment on the level of receivers' transfers. This result is not interpretable in causal terms because, as clarified in 2.2.4, lower levels of trust and trustworthiness may cause a lower success rate in the labor market; hence, the recruitment of unemployed individuals is likely to suffer from a problem of self-selection. Because of that, the experimental approach cannot disambiguate the exact direction of the causal nexus between subject's labour market state and trustworthy behavior.

The self-selection problem refers to the beliefs component of trust as inherited cultural traits (see subsection 2.2.2). In a neoclassical world, a fully rational agent, independently of her prior beliefs on others' reliability, would adjust her beliefs according to the data by repeatedly engaging in trust relevant interactions. This entails that any biased prior about others' trustworthiness is short-lived and should vanish in the long run: individuals learn the optimal level of trust and trustworthiness that maximizes their probability of not being unemployed. On this account, the unemployed state would be the effect of exogenous shocks to the economic system and not the result of individual characteristics related to trust. By contrast, culture consists of a system of beliefs that are not necessarily updated in a rational way. This entails that trust as a cultural trait does not necessarily adjust according to the available evidence accruing from repeated (either anonymous or non-anonymous) interactions. Because of that, people with a lower level of trust and trustworthiness may self-select in disadvantageous labor market states. The methodological implication of this account is that the recruitment of a representative sample of unemployed subjects should be restricted to those subjects who lost their job because of an event that is completely exogenous with respect to their individuals' characteristics.

Fehr et al (2003)'s result is consistent with the empirical (non-experimental) evidence of a negative effect of individuals' heterogeneity over behavioral trust and trustworthiness (Alesina and La Ferrara 2002, Costa and Khan 2003). Such consistency legitimates the use of the experimental method to identify the causal link of individuals' heterogeneity with trusting and trustworthy behavior. Anderson et al. (2006) implement a repeated trust game with random matching, where players' relative position is exogenously determined over rounds by varying their show up fee. The distribution of players' total endowment differs across treatments (i.e. egalitarian, symmetric and skewed) and can be either private or public to examine how the degree and forms of individuals' heterogeneity impact on behavioral trust/trustworthiness. They find, however, a weak and non-systematic negative effect of individuals' heterogeneity on trusting and trustworthy behavior. In a further study, Greiner et al.

(2012) implement a dynamic trust game, where players' income gained from the interactions is cumulated over time. In the first treatment, subjects start from an equal position, while in the second they start from an unequal one. In both treatments, players' relative position is known. This design allows disentangling the effect of endogenous (first treatment) and exogenous (second treatment) heterogeneity. The comparative static analysis of the first rounds of the two samples shows a negative effect of exogenously determined heterogeneity: i.e. the average level of transfers in the first round is lower in the second treatment than in the first one. However, the dynamic patterns of players' offers *between* the two samples tend to converge over time: specifically, the average transfers in the first treatment tends to fall down over rounds, while it stays constant in the second one. This testifies a stronger negative effect of endogenously determined heterogeneity, interpreted as a signal of receivers' untrustworthiness. Finally, Heap et al. (2013) propose an experimental design similar to Anderson et al. (2006) with the modification of informing subjects' on the counterpart state. This allow to assess whether the effect of economic heterogeneity depends on players' *specific* relative position. They find a corrosive effect of endowment inequality on behavioral trust, particularly when the counterpart's state is made public.

Overall, experimental investigations provide some evidence of a general state-dependence of trusting and trustworthy behavior. So far, however, results lack robustness and much work has to be done to purify the effect of individuals' heterogeneity from the possibly confounding learning and strategic effects, which are typical of repeated games where the financial consequences of previous rounds' choices are known. Moreover, analyses of the interaction of individuals' heterogeneity with the multiple and non-exclusive motivational determinants of trust and trustworthiness are still lacking. Finally, economic heterogeneity may derive from both positive and negative shocks to total wealth rather than being the output of a pure redistribution. However, none of the above mentioned experiments control for the differential impact of exogenously induced heterogeneity across different economic contexts.

2.7 Conclusions

The chapter delivered a conceptualization of trust in terms of preferences and beliefs, emphasizing the relevance of the dynamics of cultural transmission, which shape non-standard preferences (i.e. other-regarding preferences) and expectations (based on social learning) responsible for the emergence of trusting behaviors in one-shot anonymous interactions. Moreover, the chapter provided a definition of trustworthiness in terms of a general disposition to reciprocity, guiding subjects' behavior in a wide class of social situations, one-shot anonymous interactions included. On these grounds, the chapter surveyed of the main experimental literature aiming at identifying the motivational determinants of cooperative behavior in the trust game. Specifically, the chapter asked what this evidence exactly falsifies of the complex deductive machinery leading to the sub-game perfect Nash-equilibrium of a zero level of net transfers between senders and receivers. To answer, the chapter provided a description of the progressive development of a compounded experimental design (i.e. survey, dictator game, trust game and lottery choice), aiming at isolating the specific characteristics of individuals' motivations responsible for the emergence of cooperation in the trust game. The general upshot was that trusting behaviours are mainly due to non-standard preferences (i.e. altruism, inequity aversion, propensity to reciprocity and betrayal aversion) and to individuals' expectations about trustees' reliability. Similarly, receivers' cooperative behavior is due exclusively to other-regarding preferences, either conditional (e.g. reciprocity) or unconditional (e.g. altruism and inequity aversion). On this account, the existence of reciprocity and unconditional other-regarding preferences legitimates senders' beliefs on receivers' reliability, so as to induce them to a trusting choice. Finally, the chapter identified the possi-

ble research line of analyzing the causal effect of the controlled variation of subjects' economic state on their trusting and trustworthy behaviour, with particular concern for the interaction of the multiple and non-exclusive motivations to trust and trustworthiness with the differential impact of individuals' economic heterogeneity across economic contexts.

2.8 References

- Adriani, F., & Sonderegger, S. (2015). Trust, trustworthiness and the consensus effect: An evolutionary approach. *European Economic Review*, 77, 102-116.
- Alesina, A., & Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature*, 53(4), 898-944.
- Alesina, A., & La Ferrara, E. (2002). Who trusts others?. *Journal of Public Economics*, 85(2), 207-234.
- Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3), 715-753.
- Anderson, L. R., Mellor, J. M., & Milyo, J. (2006). Induced heterogeneity in trust experiments. *Experimental Economics*, 9(3), 223-235.
- Andreoni, J., Harbaugh, W. T., & Vesterlund, L. (2008). Altruism in experiments. *The New Palgrave, Second Edition*, edited by Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan.
- Algan, Y., & Cahuc, P. (2010). Inherited trust and growth. *The American Economic Review* 100(5), 2060-2092.
- Arrow, K. J. (1974). *The limits of organization*. Norton, New York.
- Ashraf, N., Bohnet, I., & Piankov, N. (2006). Decomposing trust and trustworthiness. *Experimental Economics*, 9(3), 193-208.
- Benoit, J. P., & Krishna, V. (1987). Nash equilibria of finitely repeated games. *International Journal of Game Theory*, 16(3), 197-204.

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

Bigoni, M., Bortolotti, S., Casari, M., & Gambetta, D. (2017). At the root of the North-South cooperation gap in Italy: Preferences or beliefs?

Bisin, A. and Verdier, T. The economics of cultural transmission and socialization, in Benhabib, J., Bisin, A., & Jackson, M. O. (Eds.). (2010). *Handbook of Social Economics* (Vol. 1). Elsevier.

Bohnet, I., Greig, F., Herrmann, B., & Zeckhauser, R. (2008). Betrayal aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States. *The American Economic Review*, 294-310.

Bohnet, I., & Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior & Organization*, 55(4), 467-484.

Bruni, L., & Sugden, R. (2000). Moral canals: trust and social capital in the work of Hume, Smith and Genovesi. *Economics and Philosophy*, 16(1), 21-45.

Cairò, I., & Cajner, T. (2016). Human capital and unemployment dynamics: why more-educated workers enjoy greater employment stability. *The Economic Journal*.

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, 117(3), 817-869.

Costa, D. L., & Kahn, M. E. (2003). Civic engagement and community heterogeneity: An economist's perspective. *Perspectives on Politics*, 1(01), 103-111.

Cox, J. C. (2004). How to identify trust and reciprocity. *Games and Economic Behavior*, 46(2), 260-281.

Dufwenberg, M., & Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2), 268-298.

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

- Fehr, E. (2009). On the economics and biology of trust. *Journal of the European Economic Association*, 7(2-3), 235-266.
- Fehr, E., Fischbacher, U., Von Rosenblatt, B., Schupp, J., & Wagner, G. G. (2003). A nation-wide laboratory: examining trust and trustworthiness by integrating behavioral experiments into representative survey. IZA Discussion paper series, No. 715.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817-868.
- Gambetta, D. (2000). Can we trust trust?, in (D. Gambetta, ed.), *Trust: Making and Breaking Cooperative*, Basil Blackwell.
- Gauthier, D. P. (1986). *Morals by agreement*. Oxford university press.
- Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring trust. *Quarterly Journal of Economics*, 811-846.
- Greiner, B., Ockenfels, A., & Werner, P. (2012). The dynamic interplay of inequality and trust — An experimental study. *Journal of Economic Behavior & Organization*, 81(2), 355-365.
- Hardin, R. (2002). *Trust and trustworthiness*. Russell Sage Foundation.
- Hardin, R. (1993). The street-level epistemology of trust. *Politics & Society*, 21(4), 505-529.
- Heap, S. P. H., Tan, J. H., & Zizzo, D. J. (2013). Trust, inequality and the market. *Theory and Decision*, 74(3), 311-333.
- Houser, D., Schunk, D., & Winter, J. (2010). Distinguishing trust from risk: An anatomy of the investment game. *Journal of Economic Behavior & Organization*, 74(1), 72-81.
- Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5), 865-889.
- Manapat, M. L., Nowak, M. A., & Rand, D. G. (2013). Information, irrationality, and the evolution of trust. *Journal of Economic Behavior & Organization*, 90, S57-S75.

Putnam, Robert; Robert Leonardi, Raphaelle Nanetti (1993). *Making Democracy Work* (NJ, Princeton University Press).

Sapienza, P., Toldra-Simats, A., & Zingales, L. (2013). Understanding trust. *The Economic Journal*, 123(573), 1313-1332.

Smith, V. L. (1998). The two faces of Adam Smith. *Southern Economic Journal*, 2-19.

Tomasello, M. (2014). *A natural history of human thinking*. Harvard University Press.

Van Rooij, M., & Van Orden, G. (2011). It's about Space, It's about Time, Neuroeconomics and the Brain Sublime. *The Journal of Economic Perspectives*, 25(4), 31-55.

CHAPTER 3: Trust and reciprocity in youth labor markets. An experimental approach to analyzing the impact of labour market experiences on young people

3.1 Introduction

This chapter reports the results of an experiment that seeks to detect the systematic differences in behavioral trust and trustworthiness amongst individuals' in different labour market states (i.e. employed, student or NEET) as well as the effect of the *information* on the counterpart's labour market status over subjects' trusting and trustworthy behavior.

As clarified in 2.2.4, the development of the capacities of entrusting others and being trustworthy has long lasting consequences on young people's labor market, both for the success rate in the job search and for the job stability. Indeed, trust and trustworthiness induce people to engage in mutually beneficial transactions so to increase the probability of being successful in the job search. Also, trust and trustworthiness increase the cooperation rate in the job place, boosting firms' benefits and - consequently - reducing the individuals' probability of being fired. The opposite holds for people with lower levels of trust and trustworthiness, which may be severely limited in their opportunities of engaging in mutually beneficial transactions as well as in their willingness to cooperate in the job place. Moreover, it is natural to think of a two-way causal relationship between trust/trustworthiness and the labor market status, since the success in the job search as well as the job stability are likely to further increase individuals' level of trust and trustworthiness. Conversely, the experience of failure in the job search and layoff may negatively affect individuals' trusting and trustworthy behavior. This downward spiral is one factor underlying the hys-

teresis evident in youth unemployment and which consequently contributes to the social exclusion of some youngsters.

On this account, employed should exhibit a higher level of trust and trustworthiness with respect to students and NEETs. This state-contingency of trust and trustworthiness entails that the information on the labor market status may be a signal of the individual-specific level of trusting/trustworthy behavior so to influence individuals' choices in trust relevant interactions. At the same time, the knowledge of the counterparts' labor market state is likely to trigger individuals' other-regarding behavior towards more disadvantaged individuals or people in the same LM status because of within group preferences.

The proposed analysis grounds in a number of different research lines primarily from labor economics but also covering experimental economics. First, several empirical contributions based on survey data (e.g. Jones & Riddell, 2000, Brandolini et al., 2004, Battistin et al., 2007) find significant behavioral differences between the unemployed and some other forms of NEET. Second, it is well established that the negative consequences of youth unemployment and non-employment persist in terms of long run wage and employment penalties (e.g. Gregg, 2001, and Gregg and Tominey, 2005, Cockx and Picchio, 2011). Third, more recent researches also provide evidence that unemployment reduces young people's trust (Eurofound 2012). This latter effect is likely to contribute to the persistence of wage and employment penalties producing both immediate and longer-term negative consequences for young people's welfare (Caspi et al. 1993; Brook, 2005).

Evidence on social capital and youth unemployment in the existing literature is based almost exclusively on survey measures. Several papers show, however, that such measures are not strongly related to the more relevant behavioral measures derivable from laboratory experiments (e.g. Glaeser et al., 2000, Fehr et al. 2003, and

Farina et al., 2009)². Amongst these, Fehr et al. (2003) examine the impact of individual characteristics on behavioral trust/trustworthiness and find a negative correlation between the unemployment state and the trusting/trustworthy behavior of a nationally representative sample of participants in a Trust Game.

This experiment involves groups of young people (18-29) drawn from the outside of the University environment. The experimental sessions have been implemented in three Countries: Hungary (Budapest), Italy (Naples) and the UK (Oxford). Subjects' employment state is elicited through survey questions inspired to the International Labour Organization (ILO) definitions of the relevant labour market states. The core of the design consists of two one-shot trust games with random and anonymous matching: in the first game, subjects received no information on the counterpart; in the second one, players' labour market state was common knowledge. Dictator games and lottery choice problems are implemented to control respectively for altruistic preferences and attitudes towards risk that might affect trusting and trustworthy behaviour. Moreover, survey and behavioural data are combined to test whether subjects' answers to attitudinal questions on trust and reciprocity predict their behavior.

The experiment shows clear differences in behavioral trust and reciprocity according to labor market status and in the way the information about the counterparts' labor market status influences young people's choices. The trial documents that NEETs are not an homogeneous category. Specifically, unemployed NEETs (i.e. non-employed individuals searching for a job *without* success) are less trusting and trustworthy than other NEETs not searching for a job. Moreover, those in precarious employment forms – and above-all those in temporary employment – are less trustworthy in their behavior. The implication is that the increasing tendency to promote flexible employment forms is likely to have long-term negative consequences for

² The only exception to these negative results is the contribution of Sapienza et al. (2013), which, due to the specific features of their design, have identified a significant correlation of survey measures of trust to both trusting and trustworthy behavior.

young people's labour market, operating through the channel of the negative effects of precarious employment on young people's trust/trustworthiness. The experiments also show that employed players reduce the level of trust/trustworthiness when they get information on the NEET status of the counterpart. Similarly, NEETs reduce their cooperation level when they know to be playing the game with employed people. These results, together with the evidence of a lower level of trust/trustworthiness from NEETs, provides a candidate mechanism for the explanation of the persistence in the long run of wage and employment penalties for young unemployed people. Finally, the experiments provide some evidence on the relevance of solidaristic motives in determining behavior. Specifically, the information on the status of students induces Employed and NEET to increase respectively their level of trust and trustworthiness.

The chapter is organized as follows. Section 3.2 describes the experimental design. Section 3.3 reports the main results, whilst section 3.4 concludes.

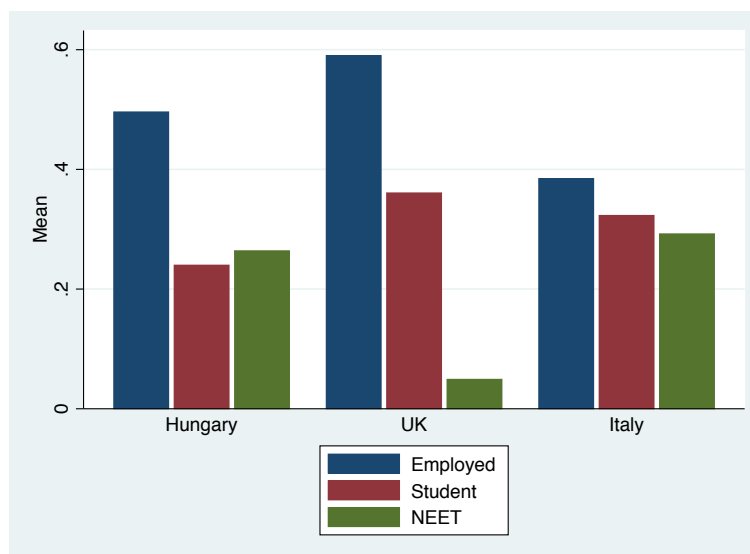
3.2 Experimental Design

The proposed approach combines more traditional survey based measures of attitudes, behavior and status with a laboratory based experiment on young people. The aim of the experimental design is to study the behavioral differences amongst individuals in different labor market states as well as the effect of information on the labor market status of their co-players on subjects' trusting and trustworthy behavior.

The entire process took place in the laboratory. Young people aged 19-29 were recruited from outside the University environment. The experiments were run in three countries: Hungary (Budapest), the UK (Oxford) and Italy (Naples). Recruitment aimed at a reasonably sized sample from each of the three broad labor market states rather than a representative sample from the youth population. As can

be seen from figure 1, this proved most difficult in Oxford where just under 10% of the sample were NEET. The overall sample comprised a little under one quarter NEET, a little over one quarter students and around one half employed. In all, the sample comprised 632 young participants, 250 in Budapest, 260 in Naples and 122 in Oxford.

Figure 1. Distribution of experimental participants (by labour market status and country)



Subjects were invited to participate through the labor office of the relevant countries in experimental sessions, which articulate in six computer-based tasks managed by a server running a z-tree script (Fischbacher 2007). Subjects received instruction phase-by-phase, so that they did not know the entire structure of the experiment at the outset. In each interactive behavioral task, subjects were randomly matched without replacement to guarantee full anonymity.

1) *Survey, part I*: At the very beginning of the experiment, subjects were required to answer a questionnaire aimed at eliciting their socio-economic characteristics.

2) *Dictator Game (DG)*: Players were randomly assigned a role of type A (sender) or type B (receiver) which they maintained throughout the experiment. Only senders were endowed with a monetary amount of 10 tokens. This endowment was

refunded to senders in all the subsequent interactive behavioral tasks. Players were then randomly (and anonymously) matched and invited to play a DG, with no information on their counterparts.

3) *Trust Game with 'no information' (TG1)*: Subjects had to play a standard trust game with no information on their counterpart. In this game, senders chose whether to give some or all of the endowment (i.e. from 0 to 10 tokens) to the receiver, which got three times the amount originally sent (i.e. from 0 to 30 tokens) and decided whether to give back some or all of it to the sender. Receivers' choices were elicited with the strategy method: i.e. the receiver declared the monetary amount she wanted to give back for any positive amount (from 1 to 10 tokens) the sender could decide to give to her. In such a way, the receiver's *actual* choice was conditional on sender's *unknown* decision.

4) *Trust Game with 'status information' (TG2)*: Players took part to a second trust game and received information on the labor market status (in Education, in Employment or NEET) of their co-respondent. This treatment allows identifying the effect - if any - of the information on the counterpart's labour market state on behavioral trust and trustworthiness. On one hand, this knowledge may affect senders' other-regarding motives towards the anonymous recipient; alternatively, senders may interpret the delivered information as a signal of recipients' expected trustworthiness. On the other hand, the information on senders' labour market state can only affect recipients' other-regarding motives in as much as they lack any strategic incentive linked to a potential financial gain.

5) *Lottery choice*: To elicit their risk preferences, players were shown a table (table 1) which lists 6 different – and increasingly risky – lotteries³. They had to choose *one* out of the six lotteries and, subsequently, a coin was tossed to decide which outcome was to be applied and players got paid with the actual outcome of the lottery at the

³ The lottery is a slight modification of those used by Eckell & Grossman (2008) and Casari et al. (2013).

end of the experiment. The expected value of the lottery outcome increases as does the distance between the better and worse outcomes from lottery 1 to 6⁴. Hence, each lottery choice provides a measure of the range of values of (Constant Relative) Risk aversion. Controlling for risk attitudes allows to see whether differences in behavioral trust observable across individuals in differing labor market states are systematically associated with differences in risk attitudes.

Table 1. Lottery choice

| | Heads | Tails |
|------------------|--------------|--------------|
| Lottery 1 | 7 | 7 |
| Lottery 2 | 9 | 6 |
| Lottery 3 | 11 | 5 |
| Lottery 4 | 13 | 4 |
| Lottery 5 | 15 | 3 |
| Lottery 6 | 17.6 | 0.4 |

6) *Survey, part II*: Players filled out a questionnaire eliciting more qualitative information on their attitudes towards trust and reciprocity, their locus of behavioral control and their family background. Information on attitudes allows examining the nature of the choices taken. This part of the survey was implemented at the end of the trial to avoid conditioning or framing responses by discussing issues related to trust, reciprocity and other relevant attitudes before the behavioral tasks.

Elements (2) – (5) involved choices which had direct financial consequences. Moreover, apart from individuals playing the role of dictators in the DG, outcomes depended on the behavior of others (TG1 and TG2) or upon chance (lottery choice).

⁴ Note however, that lottery 6 is riskier than, but has the same expected value as, lottery 5 hence allowing for the possibility of negative risk aversion (i.e. risk loving behavior).

To neutralize learning effects, players got paid only at the very end of the experiment, when, apart from dictators in the DG, they came to know of the behavior of their counterparts and hence the size of their payments.

Although cross-country differences and the related implications were not the main focus of the analysis, the three countries (Hungary, UK and Italy) were chosen to cover differing economic, institutional and cultural contexts, possibly affecting behavioral trust/trustworthiness (Alesina and Giuliano 2015). Indeed, the specific sites of the experiment (Budapest, Oxford, Naples) within the countries accentuate any such cross country differences. Hungary is a Central European country with a history of centralized planning and subsequent transition to the market with a relatively low unemployment rate (7.1% in December 2014)⁵ and a (medium level) ratio of youth to adult unemployment of 2.8 (last quarter, 2014). Budapest has a rate of aggregate unemployment significantly lower than the national average (5.2% in the last quarter of 2014). The UK is also characterized by a relatively low rate of unemployment (5.8% in the last quarter of 2014) but has a relatively high ratio of youth to adult unemployment (3.9 in the last quarter of 2014). Moreover, the UK labor market is relatively flexible with low levels of employment protection. Oxford is a relatively prosperous and well-educated part of the country with an unemployment rate significantly below the national average (3.6% in late 2014). In stark contrast, Italy has a relatively high unemployment rate – 12.4% at the end of 2014 - and a relatively high youth-adult ratio of unemployment rates (3.2 in the last quarter of 2014). Moreover, in Naples the unemployment rates are well above the national average (24.6%) with a correspondingly high rate of youth unemployment (57% in 2014). The Italian labor market is characterized by the so-called Mediterranean model with highly protected core employees (above all prime age male) and a secondary labor market for the young and – to some extent - females in precarious employment with limited access to the core. The three countries also

⁵ Data at the national level and by age are drawn from Eurostat; local unemployment statistics from national statistical services. In both cases these are labor force survey based estimates.

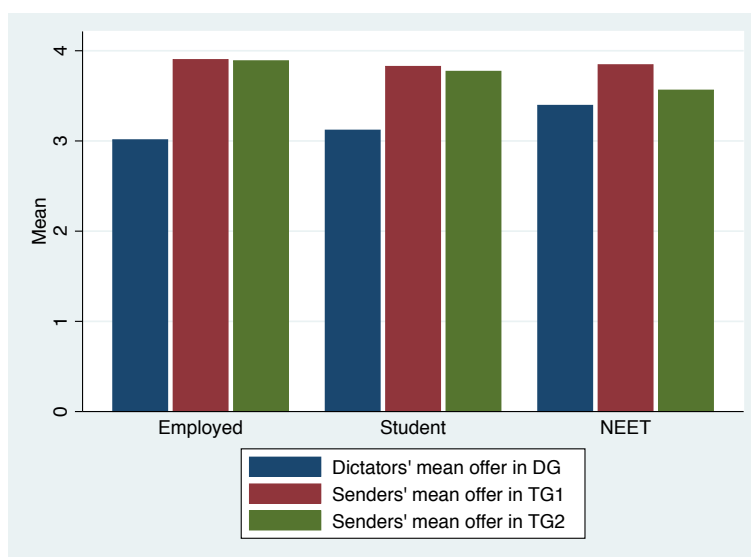
exhibit differing levels of attitudinal trust amongst its citizens, as the European Social Survey (ESS) survey 2012 testifies. Overall, Italy is the least trusting country followed by Hungary and the UK which is the most trusting. However, these differences refer to trust in institutions, rather than to trust in others, which this experiment aims at measuring behaviorally.

3.3 Results

3.3.1 Senders' behavior in TG1 (no information on the counterpart)

Figure 2 reports the behavior of senders in the Dictator Game (DG) and in the two trust games (TG1, TG2) without and with information on recipients' labor market status. Overall, senders contribute less in the dictator game than they did in the two trust games. The differences in average offers in TG1 *between* subjects in different labor market states are small and not statistically significant.

Figure 2: Amounts sent by senders in DG, TG1 and TG2 (sorted by LM status).



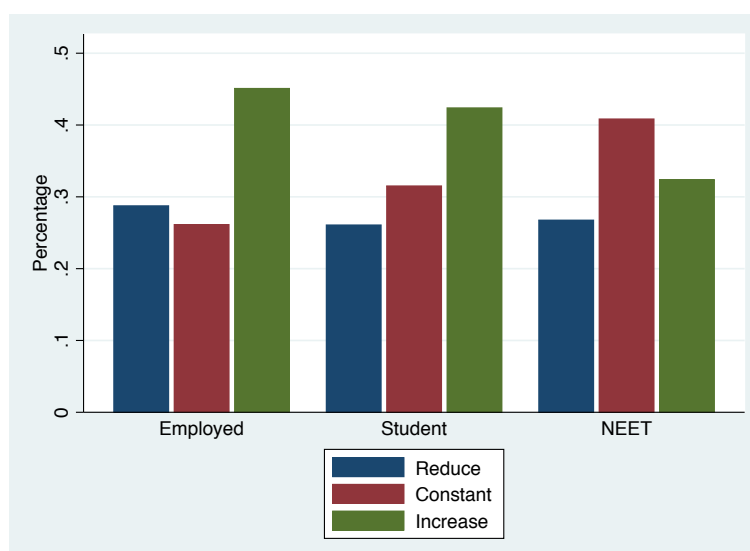
By contrast, the *within*-subjects variation in behavior between DG and TG1 is statistically significant for all senders in different labor market states. However, the difference in the average level of transfers between DG and TG1, is also relatively small - of the order of 0.8 tokens on average – suggesting that other-regarding motives are relatively strong and investment motives relatively weak in this sample. Table 2 shows the paired means comparison t-tests sorted by senders’ labor market status, while figure 3 the behavioral variations at the individual level.

Table 2. paired means comparison t-tests of senders’ giving rate between DG and TG1 (sorted by LM status)

| Employed | Student | NEET |
|----------|----------|---------|
| 0.088*** | 0.070*** | 0.045** |
| (0.000) | (0.006) | (0.033) |
| [0.265] | [0.268] | [0.204] |

* One-tailed p-values and standard deviations are in parentheses and brackets respectively.

Figure 3. Differences in senders’ behavior between DG and TG1 (by status)



Finding 1. The difference in average offers between DG and TG1 is significantly positive but decreasing in subjects' labor market status from Employed to NEET.

Table 2 documents a general tendency to significantly increase the average level of transfers between the dictator and the trust game. However, the size of the difference as well as its statistical significance is decreasing in the labor market status, from employed to NEET. Indeed, figure 3 shows that NEETs' behavioral patterns are different from those ones of the other categories of players because increasing offers is no longer the most frequent choice, whilst the majority of them (42%) choose to give exactly the same amount as in the dictator game.

Finding 2. Italians tend to reduce the average level of transfers between DG and TG1.

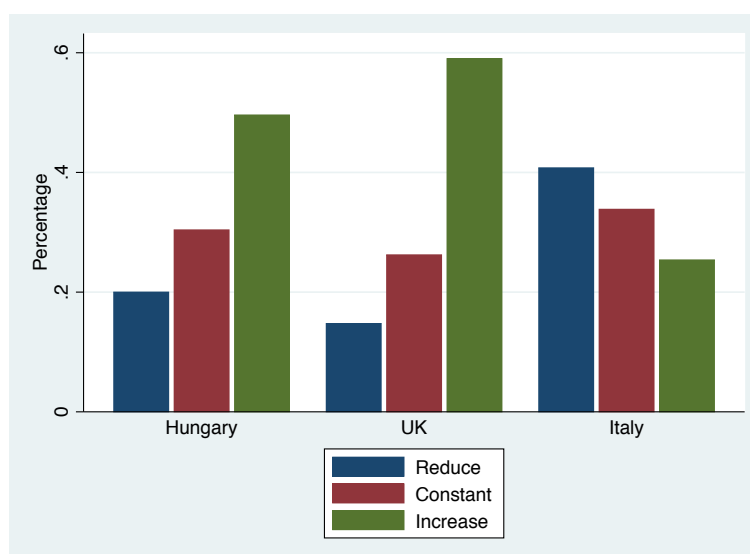
An examination of cross country differences (table 3) shows that the means difference for Italian participants is significantly negative. At the individual level (figure 4), 41% of Italian participants reduce their offers in TG1, whilst only 20% of Hungarians and 15% of British people choose to give less. This suggests that strategic motives to trust strangers are relatively strong in the UK, slightly weaker in Hungary and, at the aggregate level, entirely absent in Italy. This is broadly consistent with the levels of attitudinal trust in the countries reported by the ESS survey 2012.

Table 3. paired means comparison t-tests of senders' giving rate between DG and TG1 (sorted by country)

| Hungary | UK | Italy |
|----------|----------|----------|
| 0.107*** | 0.229*** | -0.031** |
| (0.000) | (0.000) | (0.028) |
| [0.222] | [0.331] | [0.867] |

* One-tailed p-values and standard deviations are in parentheses and brackets respectively

Figure 4. Differences in senders' behavior between DG and TG1 (by country)



Following Cox (2004), it has often been assumed that the DG is an indicator of other-regarding preferences whilst the TG provides an indication of other-regarding preferences with the addition of the 'trust' or 'investment' motive (i.e. sender's potential financial gain) which is non-negative by assumption. On this account, senders should rationally choose to give more in a TG than they do in a DG. Clearly, this condition is violated by over one quarter of the sample here, and by two fifths of the Italian participants (see figure 4). The additionality assumption, which has been questioned by Ashraf et al. (2006), does not allow for the effects of betrayal aversion (i.e. the disutility of experiencing or anticipating the non-reciprocal behavior by the

counterpart) in the decision making process (see 1.4.4). Specifically, given that receivers in the DG play a passive role, dictators do not run the risk of being “cheated on”. By contrast, in the TG the receiver can choose to reciprocate or not the sender’s trust. A failure to reciprocate may represent a non-financial cost to the sender, arising from the receiver’s ‘betrayal’ in not responding positively to her “generosity”. If this cost is sufficiently large, so as to outweigh any expected monetary return, then senders may give less in the TG than in the DG. Moreover, the subjective degree of betrayal aversion depends *inter alia* on the level of social capital in a given context. Trust in others is relatively low in Italy, and indeed, much of the early evidence on the negative effect of the lack of social capital comes from Southern Italy (Putnam et al. 1993). If, betrayal aversion is particularly strong in Southern Italy, then the trust motive is likely to be negative due to ‘betrayal aversion’ (Bohnet et al. 2009).

3.3.1.1 Trust game 1: Econometric analysis of trusting behavior

The econometric analysis here implemented allows to fine grain the identification of the determinants of trusting behavior as well as to assess the internal heterogeneity of the relevant labor market states. To this end, an ordered probit model is estimated in two specifications (table 4)⁶. First, the number of tokens senders choose to give is estimated as a function of some basic individual characteristics and variables representing labor market status. The second specification considers also variables representing attitudes and risk preferences. A term controlling for the number of tokens sent in the dictator game (called “other-regarding behavior”) is also included in both specifications. Beyond the basic

⁶ The use of the ordered probit model, though not dominant, is widespread in experimental analyses of trusting and trustworthy behavior. The main reasons of this choices refer to the discreteness and the truncated range of the dependent variable (i.e. the amount sent). See Fehr et al (2003), Johansson-Stenman et al (2011) for a methodological justification of the use of the ordered probit model for experimental analysis of behavioral trust and trustworthiness.

breakdown of participants into “Employed”, “Student” and “NEET”, the labor market status variables includes further identified forms of temporary and informal employment⁷. Within the NEETs a dummy variable was added for the unemployed referring to the sub-category of NEETs searching for a job without success⁸.

Table 4. Ordered probit estimation of the determinants of trust in TG1

| Amount sent in TG1 | coeff | SE | coeff | SE |
|------------------------------------|----------|------|----------|-------|
| Other-regarding behavior | 0.32*** | 0.01 | 0.31*** | 0.01 |
| Male | 0.11*** | 0.03 | 0.07* | 0.04 |
| Aged 25-29 | 0.17*** | 0.04 | 0.17*** | 0.04 |
| Tertiary Educ. | 0.09* | 0.05 | 0.06 | 0.05 |
| Hungary | 0.62*** | 0.00 | 0.54*** | 0.05 |
| UK | 0.88*** | 0.07 | 0.75*** | 0.07 |
| NEET | 0.03 | 0.07 | 0.02 | 0.07 |
| - & Unemployed | 0.00 | 0.07 | 0.02 | 0.07 |
| Employed | 0.01 | 0.05 | 0.03 | 0.05 |
| - & Temp. contract | -0.26*** | 0.08 | -0.26*** | 0.08 |
| - & Informal | -0.12* | 0.07 | -0.13** | 0.07 |
| Lottery (risk) | | | 0.08*** | 0.01 |
| Reciprocity | | | 0.00 | 0.006 |
| Trust | | | 0.01*** | 0.003 |
| Cooperation | | | -0.00 | 0.006 |
| Locus of Control | | | 0.01* | 0.005 |
| Pseudo R-Squared (adjusted) | 0.12 | | 0.13 | |
| n | | 316 | | |

⁷ Permanent employees were the default – excluded - category from the regression; informal employees were defined as those with no employment contract.

⁸ The sub-categories of NEET and employed were added additionally to the base (NEET or Employed) category so the effect of say being unemployed (as opposed to being a student – the excluded labor market status variable) was the sum of the coefficients on NEET and unemployed.

Consistently with previous literature (Sapienza et al. 2013, Ashraf et al. 2006) the tokens sent in the Dictator Game stand in a significant positive correlation with those sent in the Trust Game. Hence, other-regarding preferences are the crucial driver of senders' behavior in TG1. Furthermore, young men appear to be more trusting than young women as do older young people (aged over 25) with respect to younger ones⁹. Italians are the least trusting and the English the most. Although Hungarians and English young people send broadly similar amounts in the trust game, it is the English young people's behavior that changes most between DG and TG. This confirms the stronger influence of the investment motive for English people with respect to Hungarians.

Finding 3. Temporary and informal employment are negatively correlated to trusting behavior.

Being a NEET does not affect senders' behavior, while temporary employment, and, to a lesser extent informality, stand in a negative correlation with behavioral trust. The results suggest that precariousness have a more detrimental effect on trust than does non-employment per se. This is consistent with the hypothesis of lower levels of trust systematically associated to unstable forms of employment. Introducing attitudinal variables does not change the other parameter estimates greatly. Nonetheless, once risk attitudes are controlled for, the coefficient associated to the parameter on informality becomes more (not less) statistically significant ($p < 0.05$ rather than $0.05 < p < 0.10$). Hence, individuals' attitudes toward risk per se do not determine the negative correlation between precarious employment forms (informality and temporary employment) and trusting behavior.

⁹ This result is driven by the behavior of Italian young men and women. This evidence is consistent with the results reported in O'Higgins et al. (2015) which found, in contrast to analyses in other countries, that women were particularly ungenerous dictators in DG's run amongst southern Italian students; a finding which is explained in terms of the matrilineal culture of that region.

Moreover, this result confirms the tendency of risk-loving young people to choose precarious employment forms.

Turning the attention to the second specification of the model, the results suggest that risk preferences are the primary 'attitudinal' determinants of behavioral trust. Thus, the investment motive is an important factor in determining senders' trusting behavior. Consistent with the findings of Fehr et al. (2003) and Sapienza et al. (2013), the index of attitudinal trust is also positively (and statistically significantly) associated with senders' behavior, although the size of the effect is numerically smaller than that one for risk attitudes¹⁰.

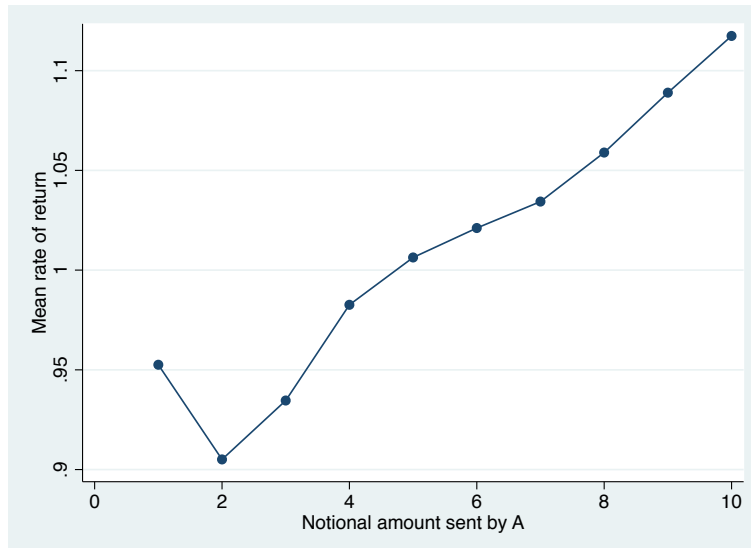
3.3.2 Trust Game 1: Receivers' behavior

The previous discussion raises the question of whether it makes sense for senders to 'trust' receivers' sense of obligation, at least in purely financial terms. The strategy method used to elicit receivers' choices is particularly well adapted to answer this question as it allows reconstructing the receivers' mean return rate - defined as the number of tokens returned by recipients divided by the number of tokens sent back - as a function of any amount the sender *may* decide to give.

On average, trusting an absolute stranger pays off; the aggregate mean rate of return is 1.01, just above the break even point of 1 (see figure 5). Moreover, consistently with the hypothesis of positive reciprocity, the return rate is increasing in the potential amount sent by the sender. Specifically, a trusting choice starts to pay off from the potential offer of an equal share of the endowment (i.e. 5 tokens).

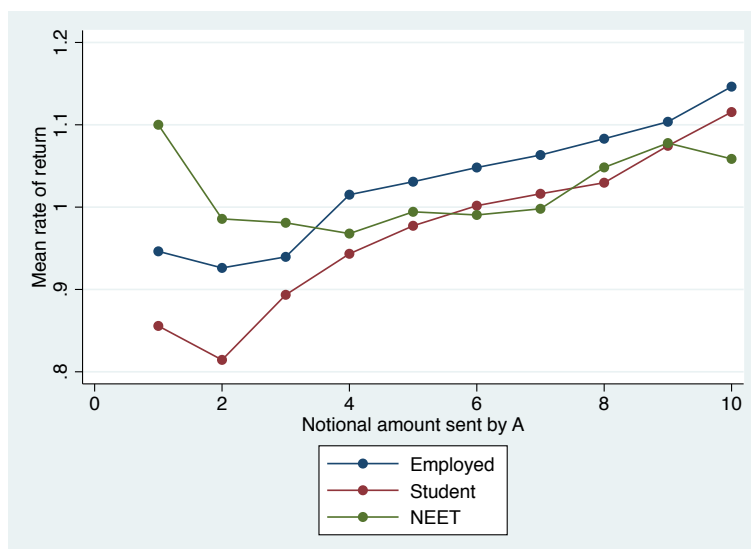
¹⁰ Trust is also positively associated with a more internal locus of control – that is with a greater sense of one's own ability to influence outcomes.

Figure 5. Rate of return in the TG1



To detail the analysis, figure 6 plots the return rates functions sorted by the labor market status. All the three categories of players behave in broadly the same way, increasing their reciprocity more than proportionately with the number of tokens the senders may give.

Figure 6. Rate of return in the TG1 (by LM status)



All three types send back 'something' when only a small amount was sent – accounting for the negative slope between one and two tokens of the return functions (see figures 5 & 6). Consistently with Sapienza et al. (2013), the negative slope between one and two tokens documents that receivers interpret senders' notional offers below the equal share of the endowment as an act of charity, which does not ask to be repaid. The employed are the most generous on average (rate of return = 1.03) and the students the least (0.97), although the latter respond the most strongly to (i.e. the rate of return increased fastest with) increases in the number of tokens sent by senders. Finally, NEETs are the least sensitive to increases in the potential amount from senders, and they guarantee a small return from the transaction only if senders offer substantially more than the half share of the endowment. Thus, on average, it is clearly profitable for senders to give half or more of their initial allocation.

3.3.2.1 Trust Game 1: Econometric analysis of trustworthiness

An ordered probit model is estimated in two specifications: i.e. without and with attitudinal variables. Due to the strategy method, ten observations per receiver are recorded; as a consequence, estimates are more precise (table 5)¹¹.

Table 5 Ordered probit estimates of the amounts sent back by B players in TG1

| Amount sent back in TG1 | coeff | SE | coeff | SE |
|-----------------------------|----------|------|----------|-------|
| sent by player A | 0.30*** | 0.01 | 0.31*** | 0.01 |
| Male | 0.10*** | 0.03 | 0.09 | 0.03 |
| Aged 25-29 | -0.04 | 0.04 | -0.04 | 0.04 |
| Tertiary Educ. | -0.15** | 0.05 | -0.12** | 0.05 |
| Hungary | 0.12** | 0.04 | 0.07* | 0.05 |
| UK | -0.27*** | 0.06 | -0.38*** | 0.06 |
| NEET | 0.30** | 0.09 | 0.38*** | 0.09 |
| - & Unemployed | -0.31*** | 0.10 | -0.39*** | 0.10 |
| Employed | 0.13** | 0.05 | 0.10** | 0.05 |
| - & Temp. contract | -0.38*** | 0.06 | -0.38*** | 0.07 |
| - & Informal | -0.06 | 0.06 | -0.03 | 0.06 |
| Lottery (risk) | | | -0.01 | 0.01 |
| Reciprocity | | | 0.002 | 0.007 |
| Trust | | | 0.02*** | 0.006 |
| Cooperation | | | 0.02*** | 0.006 |
| Locus of Control | | | 0.02*** | 0.005 |
| Pseudo R-Squared (adjusted) | 0.10 | | 0.11 | |
| N | 3160 | | | |

¹¹ Although due account is taken of the panel nature of the data in the calculation of the standard errors. To control for the non-independence of the 10 observation per individual, an ordered probit model clustering at the subject level (available upon request) has been estimated. Beside a slight reduction in precision, results did not change substantially.

As expected, risk does not play a part in the receivers' decision on the level of transfers. Otherwise, with some minor exceptions, the results are rather similar to those for trust, particularly for labor market status. Individuals on temporary contracts send back significantly less. This mirrors precisely the results for A players reported in table 4 above. A possible cognitive mechanism explaining this evidence is projection onto other individuals. Specifically, decisions by senders concerning how much to give crucially depend on how they expect recipients to react; one possible way of informing this decision is introspection – that is, how senders themselves would react in this situation (see sub-section 2.2.2)¹². If people in precarious (temporary and/or informal) employment perceive themselves to be less reciprocal on average than others, then they would exhibit a lower level of trust because they expect receivers in precarious employment to react as they would¹³.

Finding 4. NEETs not searching for a job are particularly trustworthy, while unemployed NEETs (i.e. NEET searching for a job without success) are not.

A stark contrast in trustworthiness emerges between NEETs who are not (ILO) unemployed – that is, young people without work but that do not search for it - and those who are. Individuals outside the traditionally defined labour force are particularly trustworthy (compared to the default group of student), whereas unemployed young people are not; demonstrating essentially the same degree of trustworthiness as students (i.e. the sum of the NEET and unemployed coefficients reported in the table is practically zero). Hence, the experience of failure in the job search mediates the detrimental effect of the unemployment state on subjects' cooperative behavior. Intuitively, the failure in the job search means that an

¹² See, for example, Sapienza et al. (2013) who provide explicit evidence that senders base their expectations of receivers' reactions on their own trustworthiness.

¹³ It is also arguable in this context that if it produces a positive return, as in Hungary, trusting behavior is in a (more general) sense 'rational'.

individual is believed to be untrustworthy for a given task; this may lead a person to deem herself as untrustworthy (i.e. a person not deserving trust), so to exhibit systematically a lower level of reciprocity.

The evidence of employed being more trustworthy than students is consistent with Fehr and List (2004)'s result of students exhibiting a systematically lower level of reciprocity than Chief Executive Officers. This may indicate that the influence of (conditional) other-regarding preferences is stronger for employed people than for students, because the costs of reciprocity *in the lab* are lower for people with a stable income. Alternatively, employed people may know much better than students the benefits of cooperation in the work place, so to exhibit a higher level of reciprocity in one-shot anonymous interactions.

The introduction of attitudes into the analysis again does not change the original parameter estimates very much. Consistently with Fehr et al. (2003), attitudinal trust, cooperativeness and locus of control, but not attitudinal reciprocity, are positively correlated to behavioral reciprocity.

3.3.4 Trust Game 2: The impact of information on recipients

In the second Trust Game (TG2), players' labor market state is made public: i.e. Senders know whether the person with whom they are interacting is employed, a student or a NEET. A simple way of representing the effects of this information is to examine the change in individuals' behavior between TG1 and TG2. Figure 7 reports the behavioral variations at the individual level sorted by senders' labor market status. This provides a simple benchmark to judge their interaction with the labour market status of recipients. Overall, senders tend most frequently to hold constant their offers in between the two games. However, NEETs are the most (least) likely to reduce (increase) their offers when they get informed on the counterpart's labor market status.

Figure 7. Differences in senders' behavior between TG1 and TG2, by status of senders

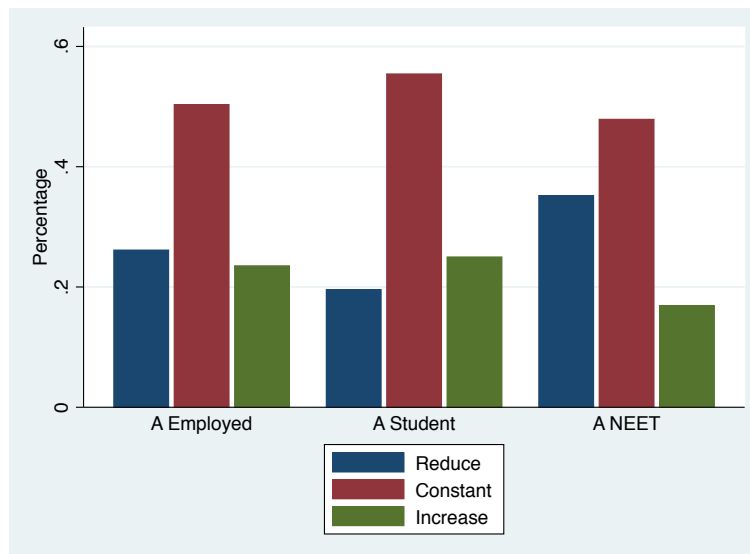


Figure 8 reports the aggregate reaction of senders to the status of recipients. Knowing that the recipient is either a student or unemployed is most likely to induce senders to increase the transfers level, whilst knowing that the recipient is employed is more likely to be associated with a fall in trust. Given the results reported above in table 4, this is more consistent with some form of other-regarding 'solidaristic' preferences rather than with the investment motive.

Figure 8. Differences in senders' behavior between TG1 and TG2, by status of receivers

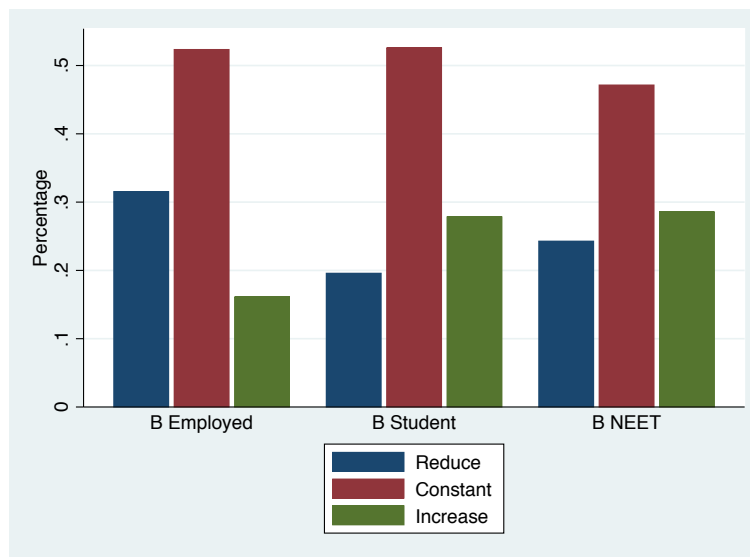


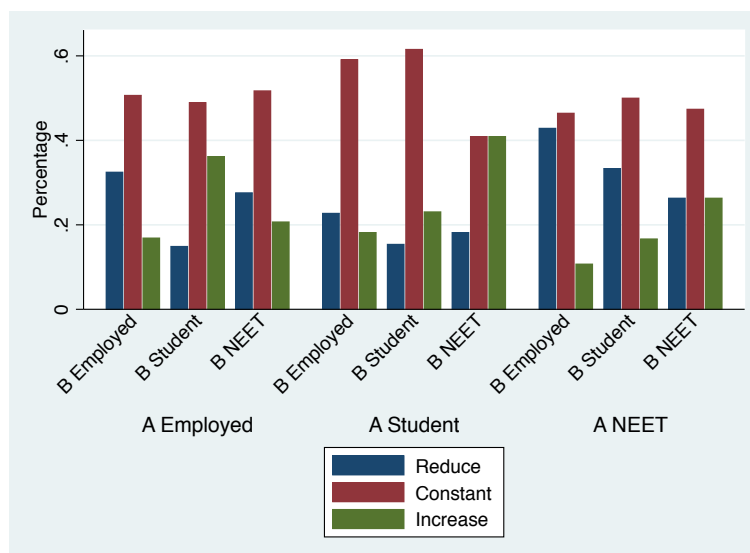
Table 6 and figure 9 tackle the issue of the effect of the interaction between the status of senders and receivers on trusting behavior. Aggregating across all states (both senders and receivers), the public knowledge of the counterpart's labor market status does not have any effect on behavioral trust. The provision of the information on receivers' LM status has a negative effect (albeit small) only on NEETs' trusting behavior. Also, all senders tend to decrease the average giving rate when they are matched with employed (table 6).

Table 6. paired means comparison t-tests of senders' giving rate between TG1 and TG2 (sorted by senders' and receivers' LM status)

| <i>Differences in average giving rate</i> | All States | B Employed | B Student | B NEET |
|---|--------------------------------|--------------------------------|-------------------------------|-------------------------------|
| All States | -0.008 (0.182) [0.167] | -0.020* (0.078) [0.178] | 0.006 (0.349) [0.156] | -0.002 (0.439) [0.156] |
| A Employed | -0.001 (0.467) [0.197] | -0.012 (0.301) [0.218] | 0.046** (0.033) [0.170] | -0.048* (0.064) [0.166] |
| A Student | -0.005 (0.353) [0.138] | -0.015 (0.202) [0.125] | -0.026 (0.173) [0.142] | 0.040 (0.112) [0.153] |
| A NEET | -0.028** (0.032) [0.126] | -0.050** (0.018) [0.120] | -0.037* (0.076) [0.124] | 0.015 (0.307) [0.134] |

* One-tailed p-values and standard deviations are in parentheses and brackets respectively.

Figure 9. Differences in senders' behavior between TG1 and TG2 (sorted by senders' and receivers' status)



Finding 5. Employed senders increase the average giving rate when matched with students and reduce it when they interact with NEET.

The means differences of the average level of transfers are significantly positive when employed senders are matched with students and negative when they interact with NEETs (table 6). Accordingly, at the individual level employed senders are more likely to increase their offers when matched with students than with NEETs (figure 9). Hence, employed senders seem to interpret the status of NEET as a signal of relative lower levels of trustworthiness. By contrast, the information on students, though revealing a category of people without an income, seems to trigger employed senders' cooperative behavior. Finally, employed do not reveal in-group preferences as the information on the employed status of the receiver does not have a significant impact on their trusting behavior (table 6). At the individual level, employed are most likely to reduce their offers when they know to be interacting with receivers of the same state (figure 9).

Finding 6. NEETs reduce the level of transfers when matched with receivers in different labor market status.

The means differences of NEETs' giving rate are significantly negative when they play the game with employed and students (table 6), with due differences for the behavioral patterns observed at the individual level. Indeed, NEETs choose more (less) frequently to increase (reduce) the giving rate when matched with students than with employed (figure 9). Moreover, NEETs increase - though not significantly - their average offer when they know to be interacting with receivers of the same status (table 6); In this case, they are more likely to increase the giving rate (figure 9). Plausibly, this is consistent with the hypothesis of in-group preferences that induce

NEETs to be relatively more other-regarding only with receivers in the same labor market status.

3.3.4.1 Econometric analysis of changes in trusting behavior

An Ordered probit model is estimated to identify the determinants of the differences in behaviour between the two trust games and *across* the nine possible matchings in TG2. In the spirit of a time differenced model, all of the attitudinal variables and preferences characteristics are not considered because they are not supposed to change through time, while the nine possible matchings between employed, students and NEETs are introduced as explanatory variables. In the proposed model, the matching Employed/Employed is the default dropped variable.

Table 7. Estimation of the determinants of changes in trust between TG1 and TG2

| <i>Difference in the average level of transfers between TG1 & TG2</i> | Pooled | |
|---|---------|-------|
| | coeff | SE |
| <i>Matching by LM status</i> | | |
| Employed/Student | 0.485** | 0.20 |
| Employed/NEET | -0.076 | 0.23 |
| Student/Employed | 0.076 | 0.19 |
| Student/Student | 0.103 | 0.23 |
| Student/NEET | 0.498** | 0.27 |
| NEET/Employed | -0.240 | 0.21 |
| NEET/Student | -0.089 | 0.23 |
| NEET/NEET | 0.249 | 0.27 |
| R-Squared (adjusted) | | 0.014 |
| n | 316 | |

Note: XXX/YYY in the table indicates that the sender was type XXX and the receiver was type YYY and, in TG2 this information is public.

Employed senders increase significantly the transfers level when they know to be matched with students compared to the default matching; i.e. employed entrust more students than other employed as can be seen from the significantly positive coefficient associated to the matching Employed/Student. Moreover, the small and non significant coefficient associated to the matching Employed/NEET indicates that employed senders exhibit approximately the same level of trust with NEETs and receivers of the same LM status.

The comparison of the coefficients associated to the matchings Student/Employed and Student/Student documents that students exhibit approximately the same level of trust with receivers of the same status and employed. Moreover, the significantly positive coefficient associated to the matching Student/NEET documents that Students exhibit a higher level of trust with NEETs compared to the default matching. By implication, students are more generous with NEET than with the other categories of players, even though the differences of coefficient is not statistically significant.

Consistently with the previous finding, NEETs exhibit a lower level of trust with employed and students compared with the default matching. Moreover, the positive (albeit not significant) coefficient associated to the matching NEET/NEET indicates that unemployed senders are more generous with receivers of the same LM status compared to employed matched with other employed. This entails that NEETs tend to entrust more group members than employed and students, as the significantly positive difference between the relevant coefficients documents.

3.3.5 The impact of information on senders

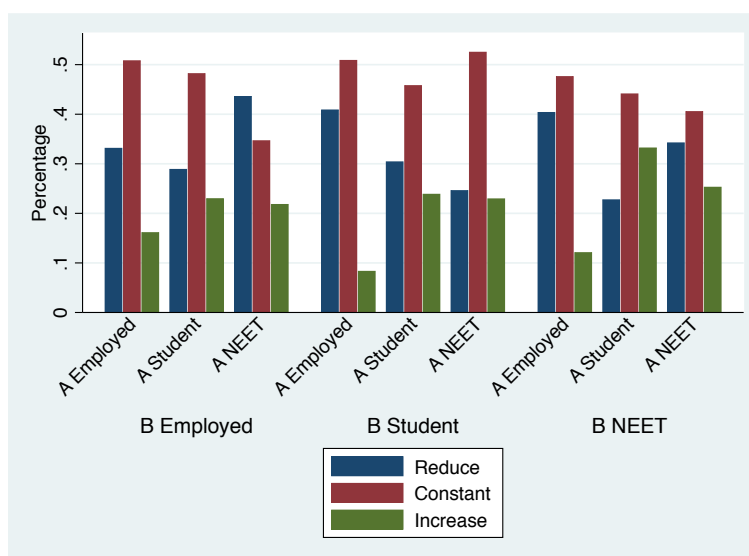
The knowledge of senders' labor market status can be informative of their willingness to trust, so to influence receivers' level of reciprocity. Table 8 and Figure 10 illustrate respectively the average and individual variations in receivers' rate of return between the two games.

Table 8. Paired means comparison t-tests of receivers' return rate between TG1 and TG2 (sorted by senders' and receivers' LM status)

| <i>Differences in average return rate</i> | All States | A Employed | A Student | A NEET |
|---|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| All States | −0.050*** (0.001) [0.973] | −0.090*** (0.000) [0.912] | 0.026 (0.749) [1.194] | −0.062** (0.013) [0.753] |
| B Employed | −0.019 (0.254) [1.155] | −0.035 (0.179) [1.081] | 0.078 (0.113) [1.356] | −0.129** (0.015) [0.993] |
| B Student | −0.116*** (0.000) [0.702] | −0.124*** (0.000) [0.802] | −0.174*** (0.000) [0.679] | −0.040*** (0.094) [0.477] |
| B NEET | −0.022 (0.243) [0.864] | −0.181*** (0.000) [0.484] | 0.160** (0.034) [0.129] | 0.006 (0.438) [0.611] |

* One-tailed p-values and standard deviations are in parentheses and brackets respectively

Figure 10. Receivers' behavioral variations between TG1 and TG2 (by receivers' and senders' LM status)



At the aggregate level, the provision of the information on the counterpart's LM status reduces the average level of trustworthiness, with a significant negative effect of the information on the employed and NEET status of the sender. Moreover, students are the most sensitive (in a negative sense) to the information on the counterpart's labor market state, whilst the information on the employed and NEET status of senders reduce the average return rate independently of receivers' state.

Finding 7. Employed receivers reduce the average return rate when matched with NEETs.

Employed receivers reduce their level of trustworthiness when they get information on the NEET status of the sender (table 8). Indeed, the matching with NEET senders markedly increases the probability of reducing the return rate from employed receivers (figure 10). The consistency of this result with employed senders' behavior (see finding 5) provides further support to the hypothesis of projection into others as

the underlying mechanism driving players behavior when the LM status is known (see 2.2.2).

Finding 8. Students reduce the average return rate when matched with employed and particularly with other students.

Students exhibit a general tendency to reduce the average level of return rate when they get information on the senders' status, but the knowledge of the employed and particularly student status of the counterpart has a stronger negative effect on their average behavior (table 8). However, the individual behavioral patterns show that being matched with an other student rises the frequency of receivers in the same state increasing the return rate with respect to the case of the matching with employed senders (figure 10). Finally, students receivers reduce less the average return rate when matched with NEETs (table 8), although at the individual level they do not significantly differ behaviorally from those students matched with other students (figure 10).

Finding 9. NEETs receivers increase the average return rate only with students.

NEET receivers significantly reduce the average level of trustworthiness when they know to be playing the game with employed senders (table 8). Indeed, they are more (less) likely to increase (reduce) the return rate when matched with students rather than with employed (figure 10). Finally, the effect of the information on the NEET status of the sender is negligible (table 8). At the individual level, a substantial part of NEETs (35%) reduce the return rate and the 25% of them (a smaller percentage with respect to NEETs matched with students) increase it.

3.3.5.1 Econometric analysis of changes in trustworthiness

The econometric analysis of receivers' behavioral variation between the two trust games and across the nine possible matchings are broadly consistent with the analysis delivered in the previous section. The results presented in table 9 follows the same logic of those ones illustrated in table 7.

Table 9. Estimation of the difference in amounts sent back by recipients

| <i>Difference in the level of trustworthiness between TG1 & TG2</i> | Pooled | |
|---|-----------|-------|
| | coeff | SE |
| <i>Matching by LM status</i> | | |
| Student/Employed | 0.114** | 0.106 |
| NEET/Employed | -0.127 | 0.08 |
| Employed/Student | -0.221*** | 0.05 |
| Student/Student | 0.245*** | 0.07 |
| NEET/Student | 0.436*** | 0.07 |
| Employed/NEET | -0.139* | 0.08 |
| Student/NEET | 0.424*** | 0.08 |
| NEET/NEET | 0.301*** | 0.10 |
| R-Squared (adjusted) | 0.011 | |
| n | 3160 | |

Source: Author calculations

Note: In this case, XXX/YYY in the table indicates that the recipient was type YYY and the sender was type XXX and, in TG2 this information is public.

Finding 10. Employed receivers are more trustworthy with students than with senders in the same LM status and NEETs.

The significant positive coefficient associated to Student/Employed documents that employed receivers are more trustworthy with students than with

senders in the same LM status. Moreover, employed are less trustworthy with NEETs than with other employed (see the significant negative coefficient associated to NEET/Employed). This entails that employed are more trustworthy with students than with NEETs, as documented by the significant negative difference between the relevant coefficients.

Students exhibit a lower level of trustworthiness with employed compared to the default matching (see the negative coefficient associated to the matching Employed/Student). Moreover, students are more trustworthy with group members than with employed, as the significant negative difference between the relevant coefficients documents. Finally, students are more trustworthy with NEETs than with the other categories, as implied by the significantly positive coefficient associated to the relevant matching.

Finding 11. NEETs are more trustworthy with players without an income.

NEETs are less trustworthy with employed senders compared with the default matching. However, NEETs exhibit a higher level of trustworthiness when they know to be matched with students and other NEETs, as testified by the positive sign and the statistical significance of the relevant coefficients. This implies that NEETs are more trustworthy with students and other NEETs than with employed; this is confirmed by the significant positive difference between the relevant coefficients.

3.4 Conclusions

The findings presented here were the result of an innovative approach that combines data on trusting/trustworthy behavior elicited *in the lab* and individuals' labor market states determined *in the real world*. Given the relatively novel approach and the number of possible aspects to study, this chapter aimed to establish the existence

of behavioral differences across labor market states, leaving for further analysis the identification of the sources of this variation.

The analysis showed that NEETs are an heterogeneous category for trusting and trustworthy behavior as has been noted by various commentators in different contexts (Furlong 2006 and Elder 2014). Contrary to the prior expectation of a negative effect of unemployment over trust and trustworthiness, NEETs, in general, are not less generous senders in the Trust Game than other young people. The econometric analysis however, allowed to qualify this basic observation. Specifically, unemployed young people (i.e. NEETs searching for a job *without* success) demonstrated much less behavioral reciprocity than NEETs who are outside the labor force as traditionally defined. Hence, the experience of failure in the job search mediates the detrimental effect of unemployment on trust and trustworthiness. A significant finding was the negative correlation between precariousness in employment and behavioral trust/trustworthiness. In recent years, several concerns have been voiced about the negative effects of the increasing prevalence of temporary employment forms on young people's early labor market experiences (O'Higgins 2010, 2012) and the results presented here appear to strongly support them. Indeed, the ongoing process of labor flexibilization is likely to have long lasting negative consequences on young people's labor market by reducing their level of trust and trustworthiness.

The analysis of the behavioral variation between TG1 and TG2 showed that knowledge of players' labor market status has a general negative effect over employed subjects' trust/trustworthiness, particularly when they know to be matched with NEETs. The latter category, in turn, exhibited a lower level of trust and trustworthiness when matched with employed players. This result suggests a candidate mechanism for the micro-level explanation of the persistence of long-term wage and employment penalties: The lower level of trust and trustworthiness of NEETs provides a rationale for employed people to distrust them; this, in turn, may

impede a successful search for a job, so to explain the persistence in the long run of wage and employment penalties to young unemployed people. Finally, the information on the student status of the counterpart induced employed and NEETs to increase the level of behavioral trust and reciprocity, documenting the existence of solidaristic motivations towards players that are not expected to provide gains from the transaction.

3.5 References

- Alesina, Alberto, Paola Giuliano. (2015). "Culture and Institutions," *Journal of Economic Literature* 53(4): 898-944.
- Arrow, K. J. (1972). Gifts and exchanges. *Philosophy & Public Affairs*, 343-362.
- Ashraf, N., Bohnet, I., Piankov, N. (2006). "Decomposing trust and trustworthiness." *Experimental Economics*, 9(3): 193-208.
- Battistin, Erich, Enrico Rettore, Ugo Trivellato. (2007). "Choosing Between Alternative Classification Criteria to Measure the Labour Force State". *Journal of the Royal Statistical Society. Series A (statistics in Society)* 170(1): 5–27.
- Berg, Joyce, John Dickhaut, Kevin McCabe. (1995). "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10(1): 122-142.
- Bohnet, Iris, Fiona Greig, Benedikt Herrmann, Richard Zeckhauser. (2008). "Betrayal Aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States". *American Economic Review* 98(1): 294–310.
- Brandolini, A., Cipollone, P. and Viviano, E. (2004) "Does the ILO definition capture all unemployment?" Temi di Discussione no. 529, Banca d'Italia, Rome.
- Brook, K. (2005). Labour market participation: the influence of social capital. Office for National Statistics, *Labour Market Trends*, March.

- Caspi, A., Wright, B.R.E.; Moffitt, T.E.; Silva, P.A. (1998). Early Failure in the Labor Market: Childhood and Adolescent Predictors of Unemployment in the Transition to Adulthood, *American Sociological Review*, 63(3), 424-451.
- Clark, K.B., Summers, L.H. Holt, C.C., Hall, R.E. and Baily, M.N. (1979). Labor Market Dynamics and Unemployment: A Reconsideration," *Brookings Papers on Economic Activity* 1: 13-72.
- Cockx, B. and Picchio, M. (2011). "Scarring effects of remaining unemployed for long-term unemployed school-leavers," Institute de Recherches Économiques et Sociale de L'Université catholique de Louvain, Discussion Paper 2011-32.
- Cox, J. C. (2004). How to identify trust and reciprocity. *Games and Economic Behaviour*, 46(2), 260-281.
- Eckel, C. C., & Wilson, R. K. (2004). Is trust a risky decision? *Journal of Economic Behaviour & Organization*, 55(4), 447-465.
- Elder, S. (2014). *What does NEETs mean and why is the concept so easily misinterpreted?* (Geneva, ILO).
- Eurofound. (2012). *Young people not in employment, education or training: Characteristics, costs and policy responses in Europe*. (Dublin, Eurofound).
- Farina, F., O'Higgins, N. and Sbriglia, P. (2009). "Suit the action to the word, the word to the action: Eliciting Motives for trust and reciprocity by Attitudinal and Behavioural Measures," *Research in Economics*, 63: 253-265.
- Fehr, E., & List, J. A. (2004). The hidden costs and returns of incentives—trust and trustworthiness among CEOs. *Journal of the European Economic Association*, 2(5), 743-771.
- Fehr, E., Fischbacher, U., Von Rosenbladt, B., Schupp, J., & Wagner, G. G. (2003). A nation-wide laboratory: examining trust and trustworthiness by integrating behavioural experiments into representative survey. *IZA Discussion paper series*, No. 715.

- Flinn, C.J. and Heckman, J.J. (1982). "Models for the Analysis of Labor Market Dynamics," in *Advances in Econometrics*, R. Basmann and G. Rhodes, eds. JAI Press, Vol. 1 (1982a): 35—95.
- Flinn, C.J. and Heckman, J.J. (1983). "Are unemployment and out of the labor force behaviorally distinct labor force states?" *Journal of Labor Economics*, 1(1): 28-42.
- Furlong, A. (2006). "Not a very NEET solution: representing problematic labour market transitions among early school-leavers", in *Work, Employment and Society*, 20(3): 553-69.
- Glaeser E.L., Laibson, D., Scheinkman, J.A. and Soutter, C.L. (2000). "Measuring Trust", *Quarterly Journal of Economics*, 65: 811-846.
- Gregg, P. (2001). "The impact of youth unemployment on adult unemployment in the NCDS," *Economic Journal*, F626-F653.
- Gregg, P. and Tominey, E. (2005). "The wage scar from male youth unemployment," *Labour Economics* 12: 487–509.
- Guiso L., Sapienza, P., Zingales, L., 2009, Cultural Biases in Economic Exchange, *Quarterly Journal of Economics*, 124(3).
- Hardin, R. (1992). The street-level epistemology of trust. *Analyse & Kritik*, 14(2), 152-176.
- Houser, D., Schunk, D., & Winter, J. (2010). Distinguishing trust from risk: An anatomy of the investment game. *Journal of Economic Behaviour & Organization*, 74(1), 72-81.
- Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5), 865-889.
- Jones, S. R. G. and Riddell, W. C. (1999) "The measurement of unemployment: an empirical approach." *Econometrica*, 67: 147-162.
- O'Higgins, Niall. (2010). "The impact of the economic and financial crisis on youth employment: Measures for labour market recovery in the European Union, Canada and the United States. ILO: Employment Working Paper No. 70." Employment Working Paper No. 70, Geneva: ILO.

- O'Higgins, Niall. (2012). "This Time It's Different? Youth labour markets during 'The Great Recession'", *Comparative Economic Studies*, 54(2): 395-412.
- Putnam, Robert; Robert Leonardi, Raphaelle Nanetti (1993). *Making Democracy Work*. (NJ, Princeton University Press).
- Putnam, R.; Goss, K.A. (2002). Introduction. in Putnam, R. (ed.) *Democracies in flux: The evolution of social capital in contemporary society*. (Oxford, Oxford University Press).
- Sapienza, P., Toldra-Simats, A., Zingales, L. (2013). Understanding trust. *The Economic Journal*, 123(573), 1313-1332.

CHAPTER 4: State-contingent trust and trustworthiness. An experimental investigation

4.1 Introduction

This experiment aims at identifying the effect of the exogenous variation of players' state (i.e. their monetary endowment) on their trusting and trustworthy behavior across three different economic contexts, where economic heterogeneity is determined through: 1) Positive shocks to total wealth; 2) Redistribution of total wealth; 3) Negative shocks to total wealth. Also, the experiment investigates the way the information on the counterpart's state mediates subjects' behavioral trust and trustworthiness.

As clarified in chapter 1, the motivations inspiring the choice of entrusting others or to reciprocate others' trust are various and non-exclusive; moreover, consistently with an account of trust and trustworthiness as cultural traits, the explanatory relevance of these multiple motivations (relevance in explaining trusting and trustworthy behavior) may change in accordance with people's economic state. Indeed, several empirical (non-experimental) contributions (Eurofound 2012, Putnam 2000, Alesina and La Ferrara 2002, Costa and Khan 2003) show that self-reported trust and trustworthiness are particularly sensitive to indexes of population heterogeneity that may induce different behavioral regularities. This entails that people's motivations to trust or to reciprocate a trusting choice may also vary with the information on the economic state of the counterpart.

The experimental results reported in the previous chapter provide evidence of a systematic negative correlation between subjects' unemployment state and their trusting/trustworthy behavior. However, the identification of behavioral differences between subjects in different LM states does not allow for any causal inference because of problems of *reverse causality* and *unobserved heterogeneity*. Indeed, the propo-

sed experiment cannot identify the right direction of the causal link between the labour market state and individuals' behaviour. Moreover, employment state and individual level of trust can be both affected by unobserved psychological characteristics, invalidating the estimated negative correlation between unemployment and behavioral trust. For example, a particularly shy and insecure person is likely to have a competitive disadvantage in the job search as well as a systematically different attitude towards other people. In this case, the correlation between unemployment and trusting behavior would be spurious.

Recent experiments try to identify the causal link between economic heterogeneity, exogenously determined through the redistribution of monetary endowments, and subjects' trusting/trustworthy behavior. Amongst these, Anderson et al. (2006) find a very weak effect of economic heterogeneity on trust and trustworthiness. Their experiment, however, raises two main problems. First, players get informed on the financial consequences of their choices at each round. Thus senders can learn about the receivers' level of trustworthiness and adjust their trusting behavior accordingly. In Anderson et al. (2006) this learning is the main driver of behavior, possibly weakening (if not offsetting) the effect of exogenously induced heterogeneity. Second, players are informed of the distribution of endowments but not of the counterpart's endowment; this does not allow to assess whether the effect of economic heterogeneity depends on players' *specific* relative position. Heap et al. (2013) address these issues and find a corrosive effect of endowment inequality on behavioral trust, particularly when the counterpart's state is made public. However, economic heterogeneity may derive from both positive and negative shocks to total wealth rather than being the output of a pure redistribution. Nonetheless, none of the above mentioned experiments control for the differential impact of exogenously induced heterogeneity across these different economic contexts.

To fill this gap, this chapter proposes a *within-subjects* design based on the random variation of subjects' state (i.e. their monetary endowment) to identify its *causal* link with trusting and trustworthy behavior as well as the effect of information on the counterpart's state on subjects' behavior. To neutralize strategic and learning effects, subjects are matched anonymously, they are not informed on the entire structure of the experiment at the outset and get to know the financial consequences of their choices that depend on the counterpart's behavior only at the end of the session.

The core of the experimental design consists of three trust games, played in different states and informational conditions. In the first one, subjects get an equal endowment and play the game in the standard way. In the second one, a random mechanism determines a distribution of "Disadvantaged", "Average" and "Advantaged" players by varying their monetary endowments. This distribution is differently implemented in three treatment groups in such a way that the degree of heterogeneity (i.e. the differences in the monetary endowment amongst the three categories of players) is held constant, but the total "wealth" changes across the three groups. Specifically, the random variation of players' endowment determines a positive shock to total wealth in treatment 1, a pure redistribution of wealth in treatment 2 and a negative shock to total wealth in treatment 3. This source of *between-subjects* variation allows identifying the differential impact of induced heterogeneity across three different economic contexts.

In the third trust game, subjects' state is the same as in the second one, but they get informed on the counterpart's state. This treatment allows understanding whether the provision of the information on subjects' specific relative position mediates their trusting and trustworthy behavior.

The design includes two dictator games, the second of which is played with inverted roles, to elicit unconditional other-regarding preferences of the whole sample (see

2.4.2 and 2.4.3). Furthermore, a standard lottery choice problem is implemented to elicit subjects' risk attitudes (see 2.4.4).

The relevant findings confirm an aggregate negative effect of induced heterogeneity on subjects' average level of trust and trustworthiness, which is mainly due to "Disadvantaged" players (both senders and receivers) reducing the level of net transfers between the first two trust games. On the contrary, the "Advantaged" state has an ambiguous and imprecise effect on behavior. Differentiating across the three treatments, heterogeneity induced through positive shocks to total wealth (Treatment 1) does not have any effect on trusting behavior, while redistribution of total wealth (Treatment 2) and negative shocks to total wealth (Treatment 3) significantly reduce the level of trust. This difference in the effect of economic heterogeneity across the three treatments is mainly driven by the behavioral variations of "Disadvantaged" senders, which significantly and substantially reduce the level of trust when they get a monetary reduction, as it is the case in the second and third treatment.

The information on the counterpart's state induces some inconsistent choices in TG3 with respect to TG2. These inconsistencies are due to the fact that cooperation in the trust game may arise from different and non exclusive motivations. Indeed, the information on the counterpart's state may be a signal indicating the player's level of trust/trustworthiness or a trigger of other-regarding behavior. Hence, the public knowledge of players' state may induce behavioral patterns in TG3 that are inconsistent with those induced by the random variation of the endowment in TG2. Indeed, information on receivers' "Disadvantaged" or "Advantaged" state lead "Disadvantaged" senders to increase the level of trust: i.e. information on "Disadvantaged" receivers trigger senders' within-group preferences, while information on "Advantaged" receivers induces senders to increase their average level of trust because they expect to earn higher profits from the transaction. Results on receivers' behavior provide further support to this interpretation. "Disadvantaged" receivers increase

the average level of net transfers when they interact with “Disadvantaged” senders, but reduce it when matched with “Advantaged” senders: since only (conditional or unconditional) other-regarding preferences can inspire trustworthy behavior, information on the “Advantaged” state of the sender indicates only a disparity in their monetary endowment, inhibiting “Disadvantaged” receivers’ altruistic or reciprocal behavior. The latter result is consistent with the experimental evidence on the negative effect of economic inequality on reciprocal behavior (Xiao and Bicchieri 2010).

The chapter proceeds as follows. Section 4.2 describes the experimental design. Section 4.3 delivers the results of the analysis of (the within subjects behavioral variations across the five behavioral tasks) individuals’ behavioral variations within the five behavioral tasks. Section 4.4 concludes.

4.2 Experimental design

A sample of 144 students from the University of Salerno without experience in economic experiments took part in a trial run at the laboratory CATI of the Department of Sociology of the same University. The random sampling procedure balanced gender and University background. Students were randomly assigned to three experimental groups of 48 subjects, each group taking three sessions. The trial was designed as a computer-based experiment managed by a z-tree script (Fischbacher 2007).

The experimental design is articulated in three phases, involving a survey and five behavioral tasks. To avoid strategic bias, subjects did not know at the outset the number of phases as well as the kind and the number of tasks they were asked to accomplish in each phase. Hence, they were informed on the specific phase they were going to face, but received the instruction for each task *separately*. In each interactive behavioral task, subjects were randomly matched without replacement to guarantee

full anonymity. Moreover, given that the financial earnings in most of the behavioral task in phases 2-3 depended on the counterpart's choice, at the end of the session a random mechanism selected the pay-off relevant games and players got paid accordingly. This procedure neutralizes learning effects.

In PHASE 1, subjects were asked to complete a questionnaire, collecting information on their socio-economic and demographic characteristics.

In PHASE 2, subjects were randomly assigned to the role of player A (sender) or player B (receiver) and these roles remained fixed throughout the experiment. Afterwards, subjects were asked to play two dictator games (DG1 and DG2) and a trust game (TG1). In the dictator game, only senders got a monetary endowment of 25 tokens (i.e. 12.5 euros) and they could choose to give the receiver any quantity between 0 and 10 tokens. The second dictator game was played with inverted roles (i.e. Player A played as receiver and Player B as sender) and dictators got 25 tokens. Players did not know the number and the kind of interactive behavioral tasks in each phase. Moreover, receivers in DG1 did not know the level of transfers from dictators. This allowed to elicit Players B's altruism in DG2 without strategic biases. In both dictator games, receivers' expectations about the dictators' level of transfers were elicited. Afterwards, subjects played a trust game (TG1) in which *both* senders and receivers got 25 tokens. In this game, senders could choose to give any amount between 0 and 10 tokens and the receiver got *three* times the amount originally sent. Receivers' choices were elicited *simultaneously* with the strategy method: i.e. the receiver declared the monetary amount (from 0 to 30) she was willing to give back for any positive amount (from 1 to 10 tokens) the sender could decide to give to her. Hence, the receiver's *actual* choice was conditional on sender's *unknown* decision.

In PHASE 3, subjects played two trust games in different states (TG2) and informational conditions (TG3). In TG2, players had 25 tokens and a random mechanism di-

vided each of the three experimental groups in three sub-groups and changed subjects' endowment, so as to determine a distribution of "Disadvantaged", "Median" and "Advantaged" players. The way this distribution was determined differed across three treatments (table 1).

Table 1. Treatments

| Treatments | Subjects | Initial endowment | Disadvantaged | Median | Advantaged |
|---------------------------------|----------|-------------------|--------------------------|--------------------------|--------------------------|
| Positive shocks to total wealth | 48 | 25 tokens | No variation (25 tokens) | + 5 tokens (30 tokens) | + 10 tokens (35 tokens) |
| Redistribution of total wealth | 48 | 25 tokens | - 5 tokens (20 tokens) | No variation (25 tokens) | + 5 tokens (30 tokens) |
| Negative shocks to total wealth | 48 | 25 tokens | - 10 tokens (15 tokens) | - 5 tokens (20 tokens) | No variation (25 tokens) |

The degree of heterogeneity (i.e. the differences in the monetary endowments amongst players of different categories) is held constant while the endowment of each category of players changes across treatments. This allows identifying the differential impact of exogenously induced heterogeneity across three different economic contexts. Moreover, the treatments disentangle the effect of the relative position (i.e. "Disadvantaged", "Median" and "Advantaged" states determined *without* variation of subjects' endowment) from the (supposedly) additional effect of the random penalization/awards imposed on players' endowment: for example, the effect of the "Disadvantaged" relative position is measured in treatment 1, where it is determined through no variation of the endowment, whilst treatments 2 and 3 measure the (additional) effect of random penalizations.

In TG3, subjects maintained the same relative position determined in the second one, but they were informed on the counterpart's state. This allows identifying the way the knowledge of players' state (i.e. the specific relative position with respect to the counterpart) mediates behavioral trust and trustworthiness.

In PHASE 4, players took part to a *lottery choice* problem. They were shown a table (table 2) which lists 6 different – and increasingly risky – lotteries. They had to choose *one* out of the six lotteries and, subsequently, the toss of a coin determined the outcome to be applied and players got paid with the actual outcome of the lottery at the end of the experiment. The expected value of the lottery outcome increases as does the distance between the better and worse outcomes from lottery 1 to 6. Hence, each lottery choice provides a measure of the range of values of (Constant Relative) Risk aversion.

Table 2. Lottery choice

| | Heads | Tails |
|------------------|--------------|--------------|
| Lottery 1 | 7 | 7 |
| Lottery 2 | 9 | 6 |
| Lottery 3 | 11 | 5 |
| Lottery 4 | 13 | 4 |
| Lottery 5 | 15 | 3 |
| Lottery 6 | 17.6 | 0.4 |

4.3 Results

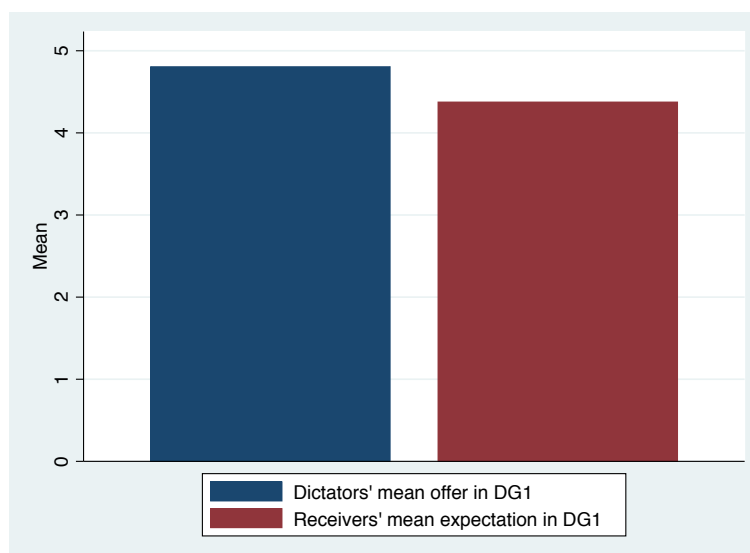
This section delivers the relevant results of the analysis of the *variation* in individuals' behavior within the five interactive behavioral tasks.

4.3.1 Dictator Game

The sample counts 72 subjects playing the role of dictators in DG1. Figure 1 shows an average offer of roughly half share of the endowment, which is significantly hi-

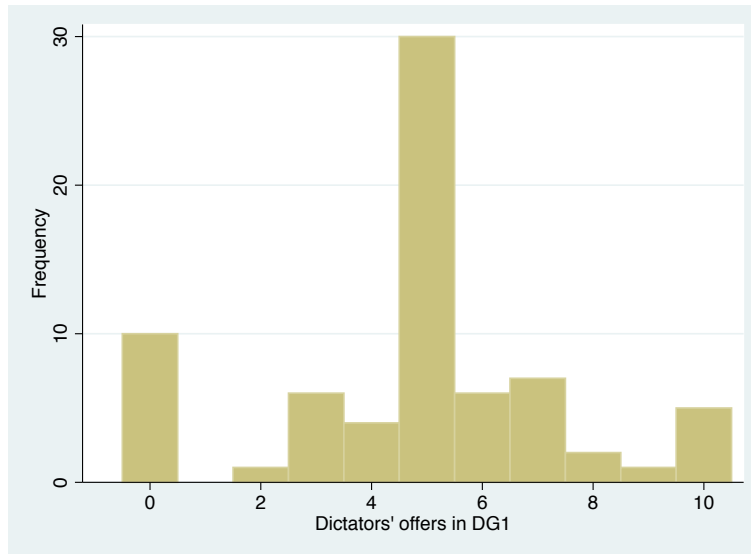
gher than the average offer of 29% identified in the experimental literature (Engel 2011). This is not, however, a source of concern given the high variance between studies and the evidence of the equal split of the endowment as one of the most frequent choice. Receivers' mean expectations are lower than dictators' actual mean offers, but the classical t-test of means comparison does not meet any conventional level of statistical significance. The distribution of offers (see figure 2) shows that the equal share of the endowment (i.e. 5 tokens) is dictators' modal choice (41%)¹⁴. However, a substantial part of them (14%) play the sub-perfect Nash-equilibrium strategy of a zero level of net transfers.

Figure 1. Average offers and expectations in DG1



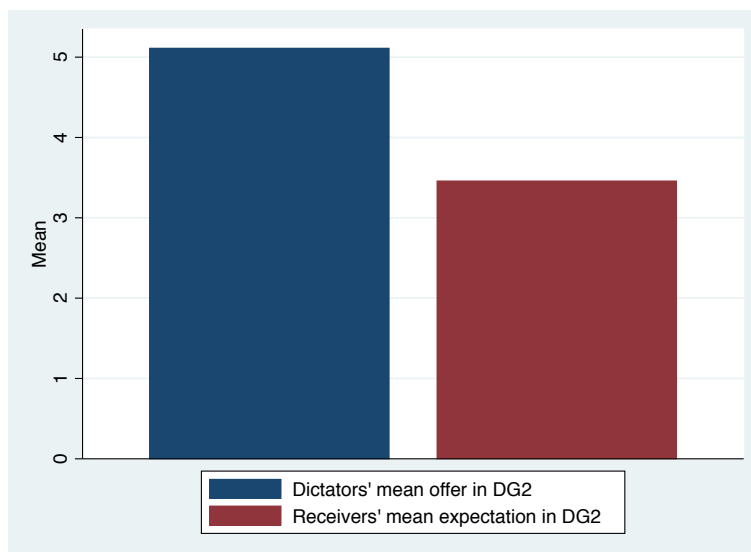
¹⁴ The distribution of receivers' beliefs (available upon request) resembles the same pattern of the offers distribution.

Figure 2. Distribution of offers in DG1



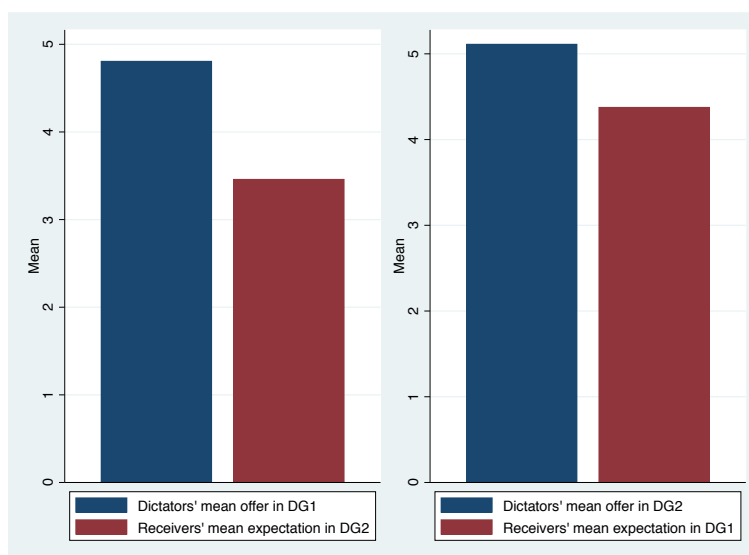
Players exhibit a similar pattern of behavior (and expectations) in the second Dictator game (Figure 3). In this case, however, there is strong evidence of a difference between dictators' offers and receivers' beliefs (Diff = 1.652 p-value = 0.00).

Figure 3. Average offers and expectations in DG2



Looking at the average offer and expectation of the *same* sample of subjects *across* the two Dictator Games (figure 4), a mismatch between players' actual offers as dictators and their beliefs as receivers is evident.

Figure 4. Dictators' offers and receivers' beliefs across DGs



Finding 1. Players' beliefs as receivers are lower than their actual choices as dictators.

In both cases, the paired-comparison t-test of means differences between players' offers as dictators and their beliefs as receivers are statistically significant (p-value = 0.00 and 0.01 respectively)¹⁵. This evidence is compatible with the hypothesis of a self-serving bias, according to which subjects believe they are more other-regarding than their counterpart; hence, they behave altruistically more than they expect other people to behave the same. Due to this self-serving bias, receivers' beliefs can be prone to order effects. If people think to be more altruistic than others, than their

¹⁵ Non-parametric tests of paired means differences (i.e. Wilcoxon signed ranks and the paired-sample sign tests) deliver the same result.

pessimism is likely to be more accentuated if they state their beliefs after having “sacrificed” some of their monetary endowment. To provide preliminary evidence to this conjecture, players’ beliefs have been regressed (with an ordered-probit model) on a dummy variable indicating the order of play of the two dictators games, together with other variables controlling for individuals’ characteristics. The significantly negative coefficient (coeff. = -0.41 p-value = 0.018) confirms that playing first as dictators and then as receivers has a negative effect over individuals’ expectations about others’ altruism with respect to the opposite order of play.

4.3.2 Differences in senders’ behavior in Dictator game 1 and Trust game 1

In the trust game, receivers can decide whether to give back to the sender a percentage of the amount she gets, from 0 to 30 tokens. Moreover, both senders and receivers get an equal endowment of 25 tokens. These game’s features are likely to alter senders’ choices with respect to the dictator game, even if in both games the sub-perfect Nash-equilibrium solution predicts a zero level of net transfers. Because of that, figure 5 compares senders’ average offer in DG1 and TG1.

Figure 5. Average offers in DG1 and TG1

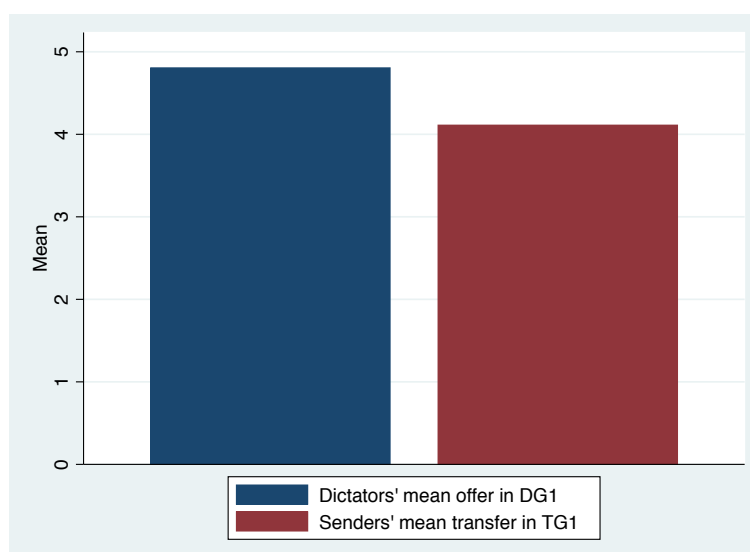
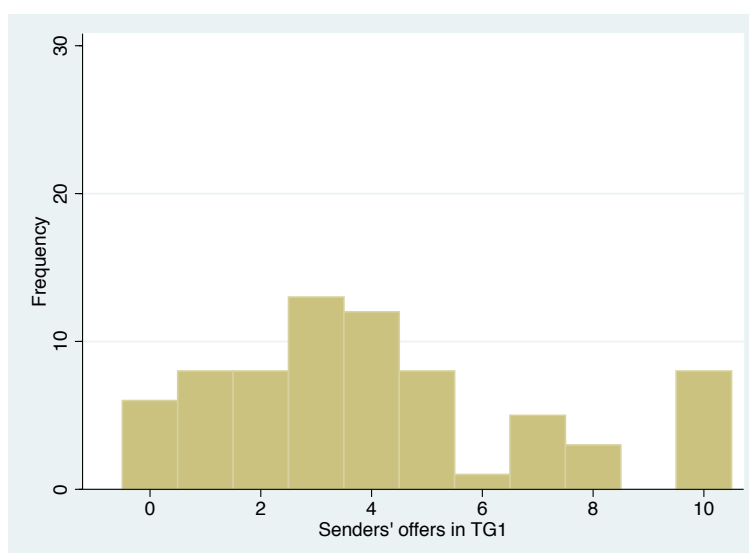


Figure 5 shows that senders offer 40% of the endowment, which is slightly lower than the average offer of 50% found in the literature (Johnson and Mislin 2011). Interestingly, senders reduce the level of net transfers in the trust game, with a mean difference significantly lower than zero (Diff = -0.694 p-value = 0.016). The distribution of offers in the first trust game (see figure 6), if confronted with the one in DG1 (figure 2), clearly shows the change in senders' behavior; indeed, the Kolmogorov-Smirnov and the Epps-Singleton tests provide strong evidence of a significant difference between the two distributions (p-value = 0.00). Moreover, the distribution of senders' offers in TG1 is significantly skewed to the right (skewness = 0.68 Pr(Skewness) = 0.01) and the equal share of the surplus is no longer the modal offer. Though players behaviors get closer to the predicted outcome of the Nash-equilibrium amongst selfish agents, a lower frequency of the Nash-equilibrium offer is observed with respect to DG1.

Figure 6. Offers' distribution in TG1

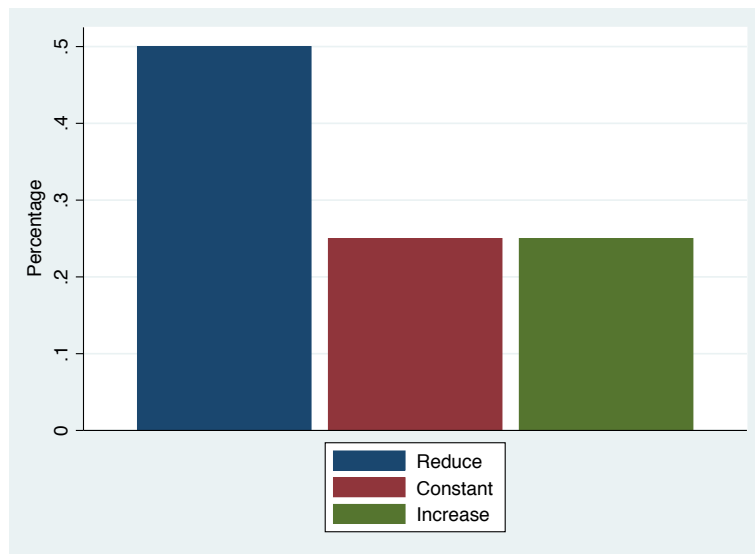


The result is anomalous with respect to the common finding of senders giving more in the trust game than in the dictator game (Ashraf et al. 2006, Cox 2004); this well-established evidence supports Cox's hypothesis of additivity of trust, according to

which the unconditional other-regarding preferences elicited in the dictator game provide the baseline level of transfer with any *positive* variation in the level of offers observed in the trust game imputed to trusting behavior. In the standard experimental setting, however, players' endowments do not vary between games, while in the proposed experiment players get an equal endowment (i.e. 25 tokens) in the trust game but not in the dictator game, where receivers do not have any monetary amount. The exogenously induced homogeneity (i.e. equality in the endowment) between players in the first trust game is likely to inhibit senders' other-regarding preferences - *particularly their concern for equity* - so as to negatively affect the average level of transfers. To test this conjecture, the amount sent in DG1 and TG1 has been regressed over a dummy variable indicating individuals' homogeneity. The significant negative result (Coeff. = -0.52 p-value = 0.003) confirms the relevance of equity concerns for the level of net transfers in the trust game.

Figure 7 tackles the issue of the individual behavioral variations between games and shows that 50% of senders reduces their offers while the remaining 50% is equally divided between senders who choose to increase or to hold them constant. Consistently with the previous literature (Ashraf 2006), the behavior of senders increasing their offers supports the hypothesis of the relevance of the investment motive in the trust game: senders increase their offers because they aim to earn profits from the cooperation with the receiver. By contrast, the behavior of senders reducing the amount sent across games is compatible with the hypothesis of inequity aversion (Fehr & Schmidt 1999): due to induced equality in TG1, senders choose to send a lower percentage of the monetary endowment they can invest in the game. Finally, senders choosing to hold offers constant exhibit a purely other-regarding behavior.

Figure 7. Behavioral variations between DG1 and TG1

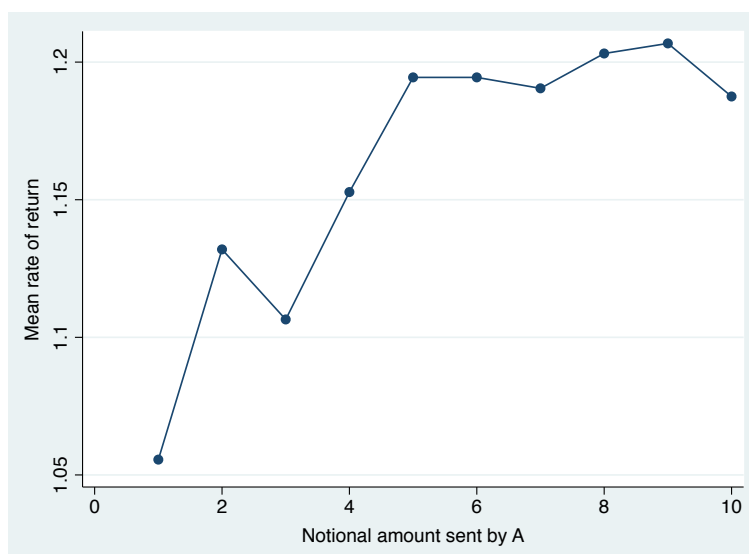


4.3.2.1 Receivers' behavior in trust game 1

Does senders' trusting behavior pay off? The strategy method is suitable to answer this question since it allows reconstructing receivers return rate (i.e. the amount sent back by receivers divided by any amount senders *may* give) as a function of the notional transfers from the sender. Figure 8 shows that the rate of return is increasing in the level of (notional) transfers from the sender. Moreover, senders' behavioral trust pays off as the function of the (average) return rate is above the break even point of 1. This result replicates the evidence of trustworthy behavior (Johnson and Mislin 2011) and supports the hypothesis of positive reciprocity that predicts higher rates of return from higher levels of (potential) transfers from A players¹⁶.

¹⁶ An apparent anomaly is the decreasing tendency of the mean return rate in between the notional offers of 2 and 3 tokens. However, the difference of the mean return rate *at these points* is not significantly different from zero.

Figure 8. Mean return rate as a function of the notional transfers from player A



4.3.3 Differences in senders' behavior between trust game 1 and trust game 2

In the second Trust Game a random mechanism reduces, holds constant or increases subjects' endowment. In each treatment, subjects are categorized as "Disadvantaged", "Median" and "Advantaged" depending on their relative position in the distribution that the random variation of their endowment determines. Subjects know their relative position and the distribution of states but not the one of their counterpart.

Table 3 shows the paired comparisons t-tests for the means differences of the average giving rate between TG1 and TG2¹⁷. At the aggregate level, senders give significantly (but not substantially) less in TG2 than in TG1. This result is consistent with the non-experimental evidence of a negative correlation between indicators of population heterogeneity and generalized trust. Moreover, differently from Anderson et al. (2006), the result shows that the negative effect of senders' heterogeneity over their

¹⁷ Non-parametric tests of means differences deliver the same results.

average level of transfer clearly emerges if the experimental protocol neutralizes strategic and learning effects.

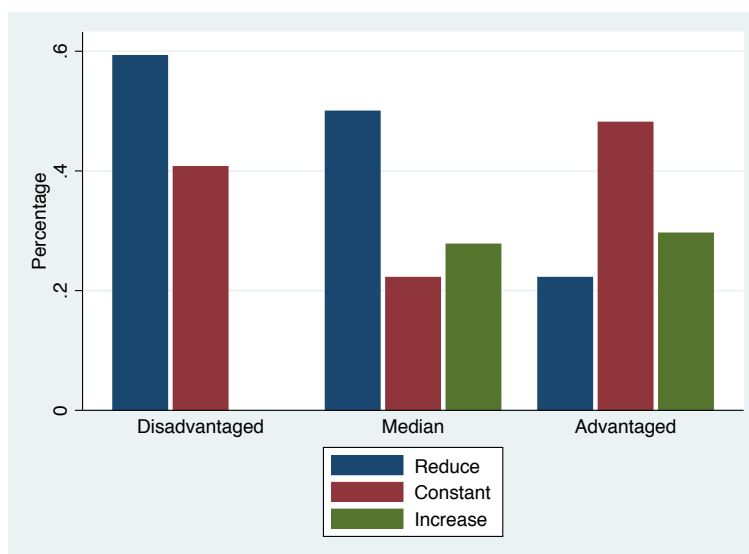
Table 3. Paired comparisons t-test of senders' giving rate between TG1 and TG2 (sorted by states)

| <i>Differences in Average Giving Rate between TG1 & TG2</i> | All Treatments |
|---|---------------------------------|
| All States | -0.061** (0.010) [0.221] |
| A "Disadvantaged" | -0.144*** (0.000) [0.155] |
| A "Median" | -0.055 (0.196) [0.268] |
| A "Advantaged" | 0.018 (0.667) [0.220] |

*One-tailed p-values and standard deviations are respectively in parentheses and brackets.

Figure 9 illustrates the behavioral variations between TG1 and TG2 at the individual level sorted by states.

Figure 9. Individual behavioral variations between TG1 and TG2 (by states)



Finding 2. Only the “Disadvantaged” state has a strong and precise negative effect on senders’ behavior.

The aggregate negative effect of players’ heterogeneity over trusting behavior is mainly driven by “Disadvantaged” senders that give significantly less on average in trust game 2 (table 3)¹⁸. Indeed, 59% of “Disadvantaged” senders reduce their offers while the remaining 41% chooses to hold them constant (figure 9). Thus, the “Disadvantaged” state induces a more selfish behavior with respect to TG1. This result echoes the finding on unemployed NEETs’ behavior in the previous experiment, confirming the tendency of the most disadvantaged categories of players to exhibit a lower level of behavioral trust (Gallego 2017).

“Median” senders give less in trust game 2, but the means difference is not statistically significant (table 3). The majority of them (50%) reduces their offers, but a consistent minority (28%) chooses to increase the level of net transfers (figure 9), so to

¹⁸ The comparison of the offers distributions of “Disadvantaged” senders in TG1 and TG2 testifies that the lower average level of offers is mainly due to the higher frequency of senders choosing the sub-perfect Nash-equilibrium offer.

make smaller and imprecise the mean negative effect of the “Median” state with respect to the “Disadvantaged” state.

Finally, “Advantaged” senders slightly increase the average level of net transfers in the second trust game but the means difference is not significantly different from zero (table 3). 48% of “Advantaged” senders hold constant their offers, but a substantial part of them (30%) increases the level of transfers and a minority (21%) chooses to reduce it. If confronted with the case of senders with a median endowment (figure 9), the imprecise positive effect of the “Advantaged” state (table 3) is mainly due to the higher (lower) percentage of senders holding constant (reducing) their offers (Figure 9).

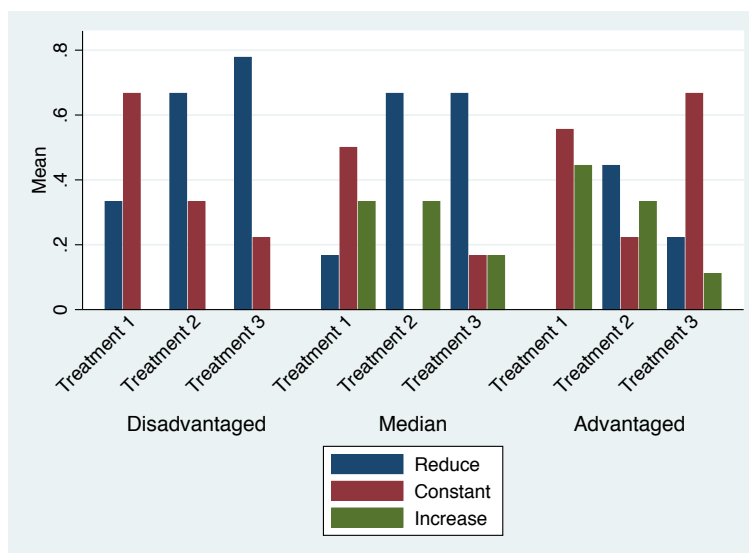
The analysis of senders’ behavior by treatments allows identifying the between subjects’ behavioral variations across the three relevant economic contexts: 1) Positive shocks to total wealth; 2) Redistribution of total wealth; 3) Negative shocks to total wealth. Moreover, given that the endowment varies across the three economic contexts for each category of players, the treatments allows disentangling the effect of the change in the relative position from the effect of the random positive/negative variations of the endowment. In this setting, the effect of the relative position is measured in those treatments where a specific state is determined *without variation* of players’ endowment. By contrast, the other treatments measure the effect of the random negative/positive variations of the endowment. Table 4 shows the paired comparisons t-tests for the means differences of the average giving rate between TG1 and TG2 sorted by senders’ state and economic contexts. The numbers in red refer to the means differences of the giving rate from players whose endowment does not change across TG1 and TG2. Figure 10 shows the individual behavioral variations between TG1 and TG2 sorted by states and economic contexts.

Table 4. Paired comparisons t-test of senders' giving rate between TG1 and TG2 (by states and treatments)

| <i>Differences in the Average Giving Rate between TG1 & TG2</i> | Positive shocks to total wealth | Redistribution of total wealth | Negative shocks to total wealth |
|---|--|---------------------------------------|--|
| All States | 0.037 (0.160) [0.181] | -0.108** (0.033) [0.276] | -0.112*** (0.001) [0.162] |
| A "Disadvantaged" | -0.078* (0.055) [0.130] | -0.155** (0.009) [0.158] | -0.200** (0.003) [0.165] |
| A "Median" | 0.100 (0.137) [0.200] | -0.133 (0.207) [0.366] | -0.133* (0.051) [0.163] |
| A "Advantaged" | 0.111** (0.042) [0.169] | -0.044 (0.344) [0.320] | -0.011 (0.380) [0.105] |

* One-tailed p-values and standard deviations are respectively in parentheses and brackets.

Figure 10. Individual behavioral variations between TG1 and TG2 (by state and treatment)



Finding 3. Heterogeneity determined through the redistribution and negative shocks to total wealth significantly reduces the average level of trust.

Aggregating across all states, wealth redistribution and negative shocks to total wealth have a significant negative impact on the mean level of transfers from sen-

ders, while positive shocks to total wealth do not have any effect. Indeed, the behavioral variations in the second and third treatment are significantly different from the ones in treatment 1 with p-values respectively of 0.018 and 0.002. Hence, the claim of an aggregate negative effect of heterogeneity on generalized trust should be restricted to those cases in which the *reduction* of the monetary endowment determines players' inequality, as it is the case in treatments 2 and 3.

The aggregate negative effect observed in the second and third treatment is mainly driven by "Disadvantaged" senders significantly reducing the average level of trust. In treatment 1, the endowment of "Disadvantaged" senders' does not vary between games and the negative effect of the disadvantaged state is relatively smaller and less precise (table 4). At the individual level, the majority of "Disadvantaged" senders (68%) hold constant their offers and a minority (32%) reduces them (figure 10). In treatment 2, "Disadvantaged" senders get a penalization of 5 tokens and lower their offers of approximately 61% with respect to the first trust game. A more precise effect is observed in treatment 3, where "Disadvantaged" senders get a penalization of 10 tokens. Moreover, "Disadvantaged" senders in treatments 2 and 3 either decrease or hold constant their offers (Figure 10). However, only the behavioral variations of "Disadvantaged" senders in treatment 3 (i.e. negative shocks to total wealth) are significantly different from the ones observed in treatment 1 (i.e. positive shocks to total wealth). Hence, "Disadvantaged" players are more sensitive (in a negative sense) to economic heterogeneity in a "poor" context than in a "richer" one.

In treatment 2 (i.e. total wealth redistribution), "Median" senders' endowment does not vary between games. In this case, the median relative position has an insignificant negative effect over trusting behavior (table 4), with players either reducing (68%) or increasing (28%) their offers (figure 10). In treatment 3 (i.e. negative shocks to total wealth), median senders get a penalization of 5 tokens and they significantly reduce their average level of transfers (table 4). Specifically, the great majority of

senders (65%) reduce their offers, while the remaining part of the sub-sample (35%) either hold constant or increase them (figure 10). In treatment 1, where the median position is determined through an award of 5 tokens, the means difference is positive (though imprecisely estimated). At the individual level, the great majority choose to hold constant (50%) or increase (32%) the transfers level, while a small minority (18%) chooses to reduce it (figure 10). Again, only the behavioral variations in treatments 1 and 3 are significantly different, confirming the stronger negative effect of poorer economic contexts on median senders' behavioral trust.

The analysis of the "Advantaged" state delivers mild evidence of a positive effect on the average offers only in treatment 3 (table 4), where senders either hold constant (56%) or increase (44%) their offers (figure 10). When the "Advantaged" state is determined through either a small award of 5 tokens (treatment 1) or without any variation of the endowment (treatment 2), it has a non significant negative effect over behavior (table 4). This confirms the asymmetric effect of the "Disadvantaged" and "Advantaged" state on trusting behavior (see finding 2).

4.3.3.1 Difference in senders' behavior between TG1 and TG2 - econometric analysis

The econometric analysis presented here pursues the objective of identifying the effect of the random positive/negative variations of senders' endowment on the variation of senders' trusting behavior between the first two games, while controlling for senders' preference characteristics (i.e. altruism and risk attitude). To this aim, an ordered probit model is estimated in two specifications (table 5)¹⁹. The first one considers only the dummies associated to positive variations of the endowment; the dummy `plus_zero` refers to "Disadvantaged" senders in treatment 1, where only positive variations of the endowment occur. The second specification considers only

¹⁹ The choice of the ordered probit model follows the convention in the literature. See Fehr et al. (2003).

negative variations of the endowment; the dummy *minus_zero* refers to the “Average” and “Advantaged” state in treatment 1 and 2 respectively where negative variations of the endowment are implemented. In both specifications the “Advantaged” state is the default (dropped) category.

Table 5. Determinants of senders’ behavioral variations between TG1 and TG2

| <i>Difference in the average level of transfers between TG1 & TG2</i> | Coefficient | Standard error | Coefficient | Standard error |
|---|-------------|----------------|-------------|----------------|
| Disadvantaged | -1.217*** | 0.375 | -0.542 | 0.368 |
| Median | -0.379 | 0.418 | -0.159 | 0.369 |
| Altruism | 0.036 | 0.051 | 0.042 | 0.057 |
| Lottery Choice (risk) | -0.098 | 0.099 | -0.090 | 0.112 |
| Plus five | 0.665* | 0.372 | . | . |
| Plus ten | 1.018*** | 0.357 | . | . |
| Plus zero | 1.101** | 0.449 | . | . |
| Minus five | . | . | -0.936** | 0.404 |
| Minus ten | . | . | -0.902** | 0.406 |
| Minus zero | . | . | -0.649 | 0.102 |
| Pseudo R-squared (adjusted) | 0.08 | | 0.07 | |
| n | 72 | | 72 | |

Table 5 shows that only in the first specification the disadvantaged state has a significant negative effect over the behavioral variations between TG1 and TG2 with respect to the missing dummy variable of “Advantaged”. By contrast, the “Median” state has a smaller and non significant negative effect in both model specifications. To control for preferences characteristics, the model includes two parameters measuring respectively individuals’ altruism (i.e. the level of transfers in DG1) and the level of risk aversion (i.e. lottery choice task). The coefficients associated to these *fixed* characteristics of individuals’ preferences are insignificantly small and they does not

affect the relevant parameters (i.e. random penalization/awards). Hence, players' altruism and risk attitude do not explain the behavioral variation of senders between the two trust games.

Finding 4. The positive and negative variations of the endowment have a symmetric effect on behavior.

The coefficients associated to the dummies of the positive/negative variations of the endowment have a significant effect of approximately the same size (in absolute terms) over the behavioral variations between the two games. Moreover, the reaction to the variation of the endowment is approximately symmetric across awards and penalizations (i.e. + 5, + 10 and - 5, - 10): i.e. Senders increase or reduce their transfers by the same amount when they are awarded or penalized of five and ten tokens.

Moreover, in Treatment 1 "Disadvantaged" senders do not get any variation of the endowment. In this case, the coefficient associated to the dummy plus_zero is significantly positive, so to mitigate the negative effect of the "Disadvantaged" relative position. This suggests that the behavioral differences observed in treatment 2 and 3 with respect to treatment 1 are mainly due to the random penalization imposed on senders' endowment.

4.3.4 Differences in receivers' behavior between trust game 1 and trust game 2

Figure 11 plots the return functions sorted by receivers' state and it documents a strong negative effect of induced heterogeneity on the profitability of trusting behavior: i.e. "Disadvantaged" and "Median" receivers exhibit a similar pattern, with a

mean return rate stably below 1, while “Advantaged” receivers guarantee some potential gain from the transaction, with a return rate stably above 1. Moreover, induced heterogeneity negatively affects the slope of the return functions, which turn out to be flatter than the return function plotted in figure 8. Since the slope of the return function measures the level of conditional cooperation, the flat shape indicates that induced heterogeneity disrupts reciprocity but not necessarily receivers’ altruism: i.e. receivers send back something mainly out of unconditional other-regarding preferences.

Figure 11. Return rate as a function of the potential transfer from senders (by state)

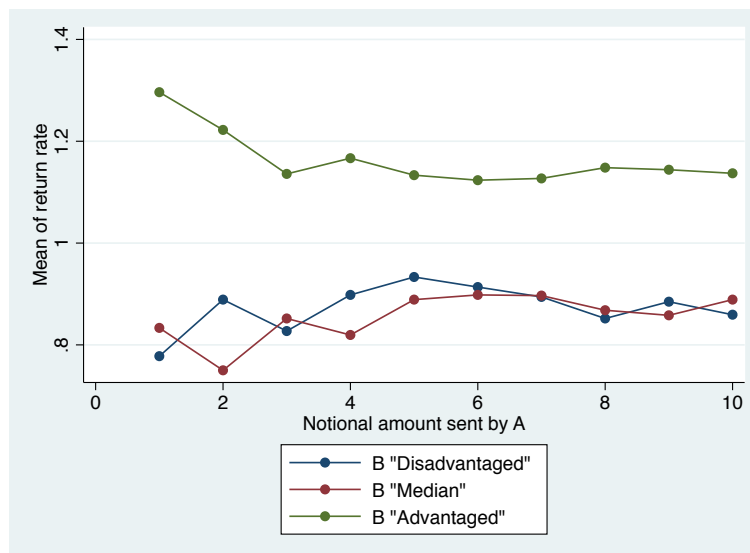


Table 6 displays the results of the paired comparisons t-test of means differences of receivers’ return rate between TG1 and TG2, while figure 12 illustrates the behavioral variations at the individual level aggregated across treatments. The strategy method delivers 10 observations for each player, so that the estimates are more precise.

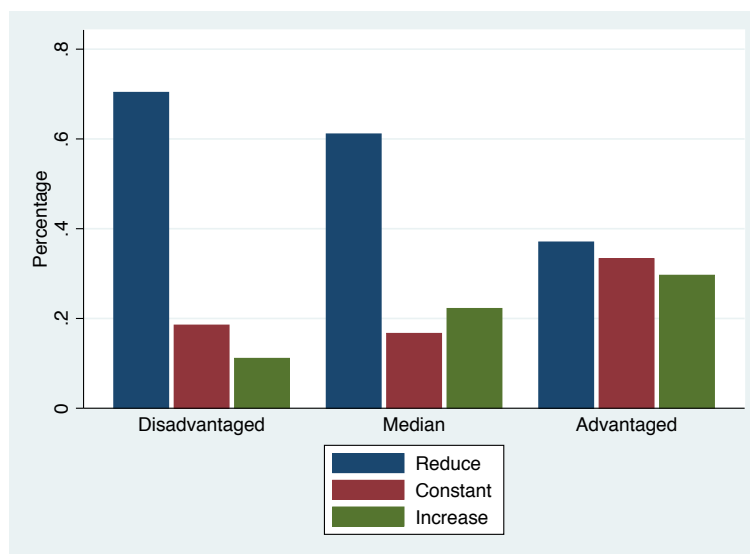
Aggregating across all states and treatments, the mean difference of the average return rate between TG1 and TG2 is significantly lower than zero. Furthermore, a general tendency to reduce the mean return rate is observed in TG2 independently of players’ state.

Table 6. Paired comparisons t-test of receivers' return rate between TG1 and TG2 (by state)

| <i>Differences in the Average Return Rate between TG1 & TG2</i> | All treatments |
|---|---------------------------------|
| All states | -0.184*** (0.000) [0.541] |
| B "Disadvantaged" | -0.317*** (0.000) [0.569] |
| B "On Average" | -0.168*** (0.000) [0.391] |
| B "Advantaged" | -0.063** (0.034) [0.569] |

*One-tailed p-values and standard deviations are respectively in parentheses and brackets.

Figure 12. Differences in receivers' return rate between TG1 & TG2 (by state)



Finding 5. Induced heterogeneity has a general negative effect, which decreases in receivers' relative position from the "Disadvantaged" to the "Advantaged" state.

The “Disadvantaged” state has a strong negative effect over the average return rate from receivers (table 6). At the individual level (figure 12), “Disadvantaged” receivers mimic “Disadvantaged” senders. Indeed, 70% of “Disadvantaged” receivers choose to reduce the return rate, while a minority of them choose to hold it constant (18%) or to increase it (12%). A smaller negative effect is observed for receivers with a median endowment (table 6). In this case, the choice of reducing the average return rate is still the most frequent one (61%), but a relatively higher percentage of receivers (22%) choose to increase it (figure 12). Finally, the “Advantaged” state has a small negative effect over receivers’ behavior (table 6). Probably this is due to the fact that “Advantaged” receivers choose to reduce, hold constant or increase the average return rate with similar frequency (figure 12). Moreover, “Advantaged” receivers are more (less) likely to increase (reduce) the individual rate of return compared to the other categories of players.

The analysis of receivers’ behavior by treatment delivers some counterintuitive result. Table 7 shows the paired comparisons t-tests for the means differences of the average return rate between TG1 and TG2 sorted by senders’ state and economic contexts. Figure 13 illustrates the individual behavioral variations between TG1 and TG2.

Finding 6. The negative effect of individuals’ heterogeneity does not vary across the three economic contexts.

Differently from the case of senders, receivers’ behavioral variations are not significantly different across treatments. In general, “Disadvantaged” receivers tend to reduce the average return rate. In treatment 2 and 3, receivers exhibit similar behavioral patterns both at the average (table 7) and at the individual level (figure 13). In treatment 1, the negative effect of the “Disadvantaged” state is smaller (table 7) and results in a lower frequency of receivers choosing to reduce the average return rate (figure 13). Consistently with the results on senders, “Disadvantaged” receivers ex-

hibit a lower level of trustworthiness in economic contexts where total wealth is either redistributed or reduced with respect to the case of positive shocks to total wealth.

“Median” receivers in treatment 2 tend to reduce the mean return rate (table 6). At the individual level, they choose most frequently to reduce the return rate (50%) but a substantial part of them (32%) choose to hold it constant or to increase it (18%). An identical negative effect is observed in treatment 3 (table 7). However, even if the choice of reducing the average return rate is still the most frequent one, a substantial part of “Median” receivers (32%) choose to increase it (figure 13). In treatment 1, the “Median” state has a significant negative effect over the mean rate of return (table 7) and at the individual level the large majority of receivers (82%) choose to reduce it. Counterintuitively, both negative and positive variations of the endowment reduce “Median” receivers’ trustworthiness (treatment 2 and 3). This result is consistent with the evidence of a negative effect of higher stakes over players’ altruistic or reciprocal behavior in decision problems in conditions of certainty: when there is more to gain, players systematically give less (Engel 2011). In treatment 1, the award of 5 tokens allows receivers to earn higher gains from the transaction so to induce them to reduce the average level of transfers.

“Advantaged” receivers do not significantly change their average behavior in treatments 2 and 3 (table 7), with some differences in individual behavior (figure 13): In treatment 2, they choose with the same frequency to reduce, hold constant or increase the return rate, while in treatment 2 they choose most frequently to hold it constant. In treatment 1, “Advantaged” receivers significantly reduce the average return rate (table 7). Indeed, they choose most frequently (58%) to reduce it, while a substantial part of them (32%) choose to increase it. Hence, a strong negative effect of the random award of ten tokens is observable in treatments 1, providing further

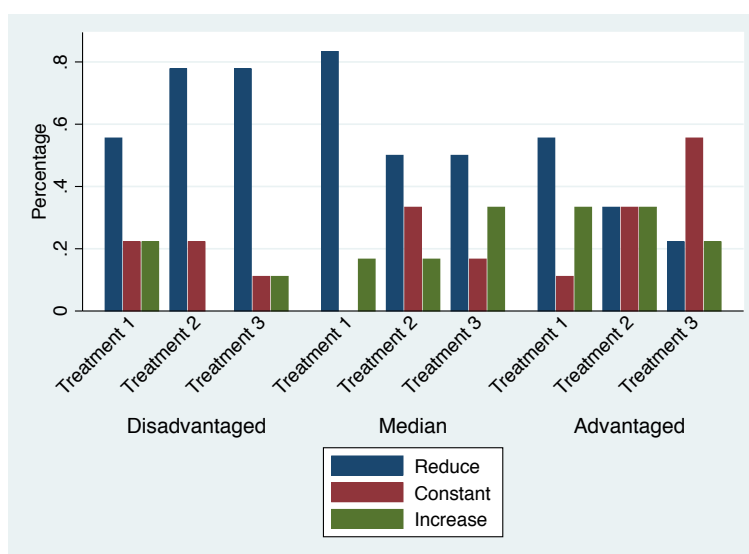
support to the hypothesis of a negative effect of higher stakes over the average level of transfers from receivers.

Table 7. Paired comparisons t-test of receivers' return rate between TG1 and TG2 (by state and treatment)

| <i>Differences in the Average Return Rate between TG1 & TG2</i> | Positive shocks to total wealth | Redistribution of total wealth | Negative shocks to total wealth |
|---|---------------------------------|---------------------------------|---------------------------------|
| All states | -0.188*** (0.000) [0.687] | -0.172*** (0.000) [0.449] | -0.193*** (0.000) [0.454] |
| B "Disadvantaged" | -0.214*** (0.002) [0.716] | -0.378*** (0.000) [0.507] | -0.359*** (0.000) [0.438] |
| B "Median" | -0.201*** (0.000) [0.432] | -0.111*** (0.005) [0.331] | -0.193*** (0.000) [0.402] |
| B "Advantaged" | -0.154** (0.033) [0.793] | -0.007 (0.424) [0.373] | -0.027 (0.279) [0.444] |

*One-tailed p-values and standard deviations are respectively in parentheses and brackets.

Figure 13. Differences in receivers' return rate between TG1 & TG2 (by state & treatments).



4.3.4.1 Difference in receivers' behavior between TG1 and TG2 - Econometric analysis

The estimation of the ordered probit model (see table 8) for the differences in receivers' behavior between the first two trust games shows that the "Disadvantaged" and "Median" states reduce receivers' level of transfers in TG2 with respect to the default category of "Advantaged" receivers. As in the case of senders, the coefficient associated to the parameter of altruism (i.e. the level of transfers in DG2) is small and non significant in both model specifications.

Moreover, the analysis identifies a significant negative effect of the notional amount sent by the sender on receivers' behavior: the more tokens the sender may give, the fewer tokens the receiver sends back to him. This provides further support to the hypothesis according which higher stakes reduce the average level of reciprocity. Consistently with this account, the coefficients associated to the dummies of the positive variations of the endowment are negative (though not significant). Indeed, the random award of 5 tokens induces to further reduce the level of transfer back when it determines a "Median" state in treatment 1 and negatively impacts over the level of transfers from "Advantaged" receivers in treatment 2. Similarly, the award of 10 tokens induces "Advantaged" receivers in treatment 1 to reduce their average level of trustworthiness.

Table 8. Determinants of receivers' behavioral variations between TG1 and TG2

| <i>Difference in the level of transfers back between TG1 & TG2</i> | Coefficient | Standard error | Coefficient | Standard error |
|--|-------------|----------------|-------------|----------------|
| Disadvantaged | -0.724*** | 0.115 | -0.456*** | 0.126 |
| Average | -0.279** | 0.108 | -0.192* | 0.109 |
| Sent by the Sender | -0.065*** | 0.013 | -0.065*** | 0.013 |
| Altruism | -0.007 | 0.013 | 0.001 | 0.013 |
| Plus five | -0.137 | 0.110 | . | . |
| Plus ten | -0.168 | 0.155 | . | . |
| Plus zero | 0.285* | 0.151 | . | . |
| Minus five | . | . | -0.213* | 0.112 |
| Minus ten | . | . | -0.172 | 0.142 |
| Minus zero | . | . | 0.068 | 0.103 |
| Pseudo R-Squared (adjusted) | 0.025 | | 0.024 | |
| n | 720 | | 720 | |

4.3.5 Differences in senders' behavior between trust game 2 and trust game 3 (information on players' state)

In the third trust game (TG3), subjects are informed on the counterpart's state. The provision of this information allows assessing whether the effect of individuals' heterogeneity depends on players' specific relative position (i.e. players' position with respect to the counterpart's position).

Table 9 shows the means differences between TG2 and TG3 sorted by the nine possible matchings, while figure 14 the individual behavioral variations. At the aggregate level, the knowledge of the counterpart's state does not determine any variation in the average transfers from senders with respect to TG2. Interestingly, the provision of information has a significant positive effect on "Disadvantaged" senders independently of the specific matching. By contrast, the knowledge of the counterpart's

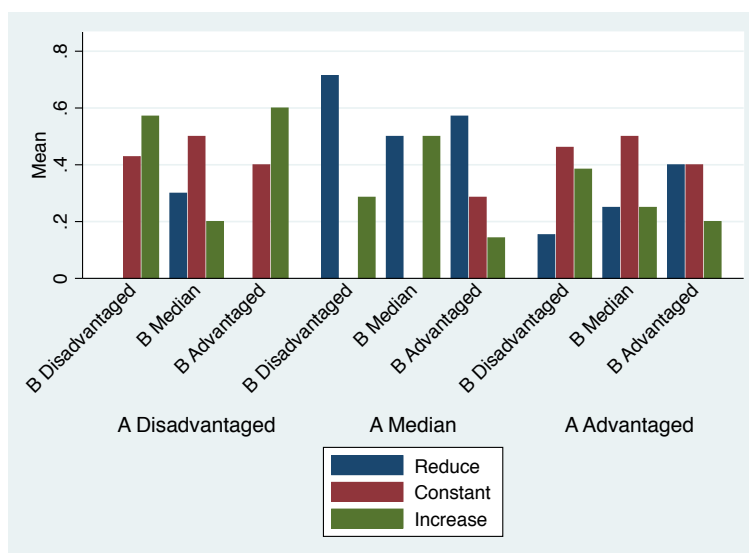
state has a general negative effect on “Median” senders and no effect at all on the “Advantaged” ones.

Table 9. Paired comparisons t-test of senders’ giving rate between TG2 and TG3

| <i>Differences in Average Giving Rate between TG2 & TG3</i> | All states | B “Disadvantaged” | B “Median” | B “Advantaged” |
|---|--------------------------------|--------------------------------|------------------------------|-------------------------------|
| All states | 0 (1.000) [0.177] | -0.0074 (0.387) [0.132] | 0.027 (0.261) [0.180] | -0.011 (0.395) [0.215] |
| A “Disadvantaged” | 0.062*** (0.008) [0.127] | 0.057** (0.015) [0.053] | 0.010 (0.398) [0.119] | 0.120** (0.018) [0.154] |
| A “Median” | -0.083* (0.095) [0.259] | -0.128** (0.046) [0.170] | 0.125 (0.252) [0.330] | -0.157* (0.083) [0.263] |
| A “Advantaged” | -0.007 (0.384) [0.129] | 0.023 (0.213) [0.101] | -0.025 (0.358) [0.125] | -0.04 (0.231) [0.164] |

*One-tailed p-values and standard deviations are respectively in parentheses and brackets.

Figure 14. Differences in senders’ behavior between TG2 & TG3 (by senders and receivers’ state)



Finding 7. Disadvantaged senders increase the average level of transfers when they interact with “Disadvantaged” and “Advantaged” receivers.

When they know to be matched with receivers in the same state, “Disadvantaged” senders either increase (58%) or hold constant (42%) the amount sent (figure 14). Interestingly, “Disadvantaged” senders significantly increase their offers when they know to be playing with “Advantaged” receivers (table 9), with an individual behavioral pattern similar to the case of matching with “Disadvantaged” receivers, but with a slightly higher probability of increasing the amount sent (figure 14). On the one hand, the matching between “Disadvantaged” players seems to trigger senders’ within-group cooperation inducing them to increase their offers. On the other hand, from the viewpoint of “Disadvantaged” senders, receivers’ “Advantaged” state increases the probability of gains from the transaction: i.e. If a norm of reciprocity drives receivers’ behavior, then a positive variation of the endowment reduces the cost of giving, so as to lead them to increase the level of transfers back. “Disadvantaged” senders may anticipate this and increase their average transfer accordingly to earn higher profit from the transaction.

Finding 8. “Median” senders reduce the average level of transfers when they interact with receivers of different states.

At the individual level (figure 14), “Median” senders either reduce (70%) or increase (30%) their offers to “Disadvantaged” receivers. Similarly, when they know to be matched with with “Advantaged” receivers, they significantly reduce the mean level of transfers, even if a lower frequency of senders reducing their offers is observed with respect to the matching with “Disadvantaged” players (figure 14). Hence, the information on the counterpart’s state makes “Median” senders more sensitive - in a negative sense - to players’ heterogeneity with respect to TG2, where the random determination of the median relative position (with the exception of treatment 3) does not have a precise impact on their behavior (see table 4).

4.3.5.1 Differences in senders' behavior between TG2 and TG3 - Econometric analysis

An Ordered probit model is estimated to identify the differential impact of the information on the counterpart's state on senders' behavioral variations across the 9 possible matchings. In table 10, the default (dropped) matching is "Disadvantaged *vs* Disadvantaged".

Table 10. Determinants of senders' behavioral variations between TG2 and TG3

| Difference in the level of transfers between TG2 & TG3 | Coefficient | Standard error |
|--|-------------|----------------|
| Disadvantaged <i>vs</i> Median | -0.477 | 0.339 |
| Disadvantaged <i>vs</i> Advantaged | 0.355 | 0.330 |
| Median <i>vs</i> Disadvantaged | -1.410** | 0.551 |
| Median <i>vs</i> Median | 0.120 | 0.953 |
| Median <i>vs</i> Advantaged | -1.368*** | 0.486 |
| Advantaged <i>vs</i> Disadvantaged | -0.295 | 0.286 |
| Advantaged <i>vs</i> Median | -0.663 | 0.463 |
| Advantaged <i>vs</i> Advantaged | -0.794** | 0.436 |
| Pseudo R-Squared | 0.073 | |
| n | 72 | |

The coefficient associated to the matching "Disadvantaged *vs* Advantaged" measures senders' behavioral variation with respect to the default matching; thus, the positive sign (though not significant) shows that "Disadvantaged" senders give more to "Advantaged" than to "Disadvantaged" receivers. This difference suggests that the investment motive induces a stronger variation in trusting behavior than other-regarding motivations towards receivers of the same category. Finally, the statistically significant difference between the coefficients associated to the matchings "Disadvantaged *vs* Median" and "Disadvantaged *vs* Advantaged" entails that "Disadvantaged" senders tend to entrust more "Advantaged" rather than "Median" receivers.

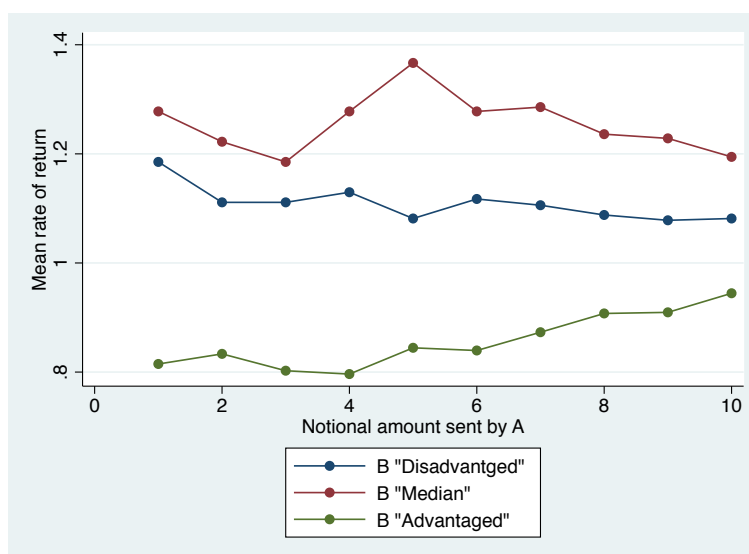
“Median” senders give significantly less to “Disadvantaged” and “Advantaged” receivers compared to the default matching, as documented by the negative coefficients associated to the relevant matchings. Furthermore, “Median” senders matched with “Median” receivers exhibit approximately the same level of trust as in the default matching (see the non-significant positive coefficient associated to “Median *vs* Median”). This entails that “Median” senders tend to entrust more receivers’ in the same state rather than other categories of players, as can be seen from the significant differences between the relevant coefficients.

“Advantaged” senders do not show significant variations with respect to the default matching, with the exception of the matching with receivers in the same state, where they exhibit a significantly lower level of trust. However, the test of the differences between the relevant coefficients does not deliver significant results.

4.3.6 Differences in receivers’ behavior between trust game 2 and trust game 3 (information on players’ state)

To start analyzing the effect of the knowledge of players’ state on receivers’ behavior, figure 15 plots the mean return rates sorted only by senders’ state. This provides a benchmark to assess the interaction effect with receivers’ state.

Figure 15. Mean return rate (by senders' state)



When receivers interact with “Disadvantaged” and “Median” senders, their mean return rate is stably above 1. By contrast, when receivers know to be matched with “Advantaged” senders, their return rate function is weakly increasing in the notional amount sent by the first player, but stably below the break-even point²⁰.

To refine the analysis, figure 16 plots the functions of the mean return rate sorted by receivers’ and senders’ state. “Disadvantaged” receivers guarantee a small return from the transaction only to “Disadvantaged” senders. When they interact with “Median” senders, the mean return rate is above 1 only for the potential offers of 1 and 5 tokens; otherwise it falls below the break even point. The return rate from “Disadvantaged” receivers interacting with “Advantaged” senders is increasing but consistently below 1. This behavioral pattern suggests the existence of “Disadvantaged” receivers’ within-group preferences.

²⁰ To put it in terms of the variation of the average return rate between TG2 and TG3, receivers matched with disadvantaged senders increase their mean rate of return. Similarly, receivers interacting with “Median” senders slightly increase it. Finally, receivers matched with “Advantaged” senders reduce the average return rate. Graphs on the variations of the mean return rate between TG2 & TG3 sorted by state are available upon request.

“Median” receivers’ return function is increasing and stably above 1 only when they interact with receivers of the same state. Furthermore, they are relatively more generous with “Disadvantaged” than with “Advantaged” senders, but both return functions are below 1. Finally, “Advantaged” receivers do not exhibit any difference in the mean return rate across senders’ state: i.e., they guarantee a small return from the transaction as the mean return rate functions are above 1 independently of senders’ relative position.

Figure 16. Return rate in TG3 (by receivers and senders’ state)

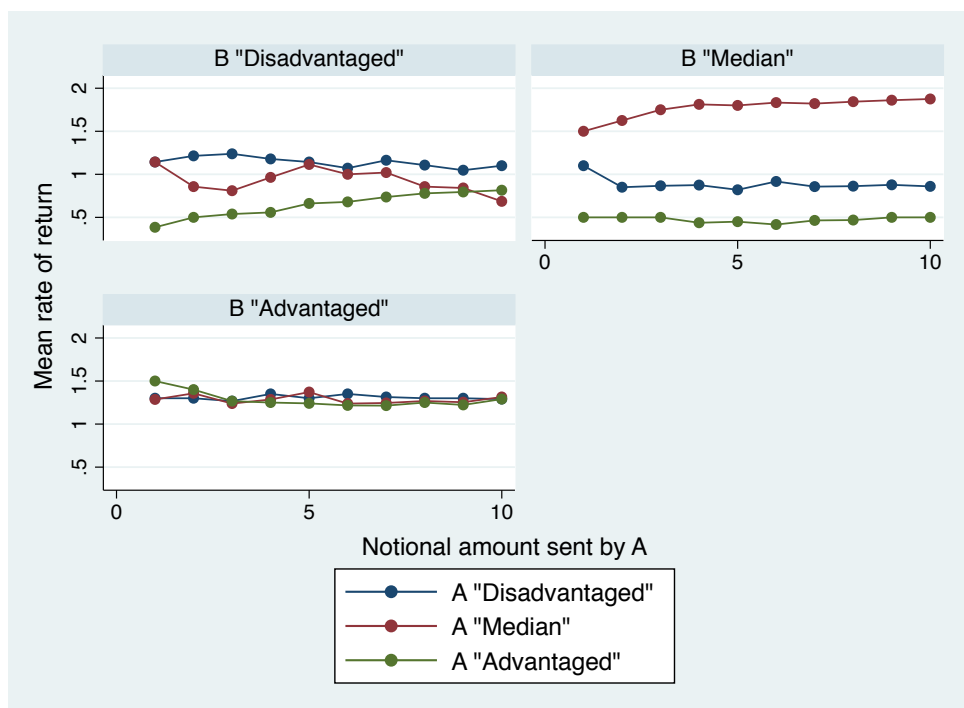


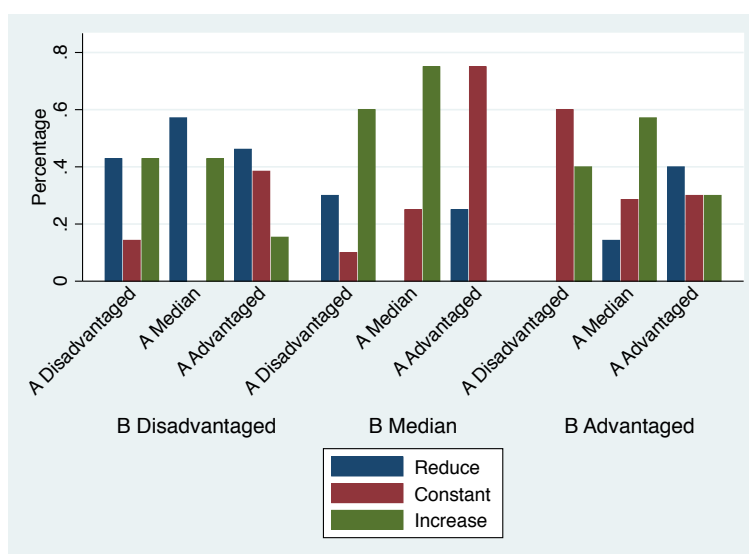
Table 11 shows the means differences of receivers’ return rate between TG2 and TG3 sorted by the 9 possible matchings, while figure 17 the related behavioral variations at the individual level.

Table 11. Paired comparisons t-test of receivers' return rate between TG2 & TG3

| <i>Differences in Average Return Rate between TG2 & TG3</i> | All states | A "Disadvantaged" | A "Median" | A "Advantaged" |
|---|--------------------------------|--------------------------------|--------------------------------|---------------------------------|
| All states | 0.073*** (0.000) [0.439] | 0.172*** (0.000) [0.434] | 0.184*** (0.000) [0.428] | -0.099*** (0.000) [0.397] |
| B "Disadvantaged" | -0.025 (0.174) [0.452] | 0.224*** (0.000) [0.538] | 0.015 (0.354) [0.342] | -0.182*** (0.000) [0.387] |
| B "Median" | 0.137*** (0.000) [0.382] | 0.046* (0.051) [0.285] | 0.517*** (0.000) [0.518] | -0.016** (0.012) [0.044] |
| B "Advantaged" | 0.129*** (0.000) [0.445] | 0.260*** (0.000) [0.451] | 0.164*** (0.000) [0.336] | -0.024 (0.297) [0.464] |

*One-tailed p-values and standard deviations are respectively in parentheses and brackets.

Figure 17. Differences in receivers' behavior between TG2 & TG3 (by senders and receivers' behavior).



The provision of information on the counterpart's state has an aggregate positive effect on receivers' behavior: independently from receivers' state, the information on the disadvantaged and median state of senders induces to increase the mean return rate, while the information on "Advantaged" senders has a negative effect on receivers' behavior. Aggregating across senders' state, "Median" and "Advantaged" recei-

vers tend to increase the mean return rate between TG2 and TG3, while the information on senders' state does not have a clear effect on "Disadvantaged" receivers.

Finding 9. The aggregate positive effect of the information on the counterpart's state is mainly driven by within group cooperation and solidarity.

Disadvantaged receivers increase the mean return rate when matched with "Disadvantaged" senders (table 11), but this positive effect masks a polarized behavior at the individual level, with 84% of "Disadvantaged" receivers either reducing (42%) or increasing (42%) the mean return rate (figure 17). The information on "Advantaged" senders has a significant negative effect on the average behavior of "Disadvantaged" receivers (Table 11). At the individual level (figure 17), the majority of individuals in this category reduces the return rate (43%), while the remaining part of them chooses to either hold constant (38%) or increase it (18%). The information on "Disadvantaged" senders seems to trigger "Disadvantaged" receivers' within group preferences, consistently with the observed behavioral patterns of senders (see tables 9 and 10). By contrast, since the profit motive is completely absent for receivers, the information on the counterpart's "Advantaged" state indicates a disparity in the monetary endowment, possibly inducing "Disadvantaged" receivers to lower the transfers level in order to reduce inequality.

Table 11 shows a small positive means difference when "Median" receivers interact with "Disadvantaged" senders, but they *substantially* (and significantly) increase the mean return rate when they know to be matched with senders of the same state. At the individual level, this latter information induces "Median" receivers to increase most frequently the return rate (figure 17). The information on senders' "Advantaged" state has a small negative effect on "Median" receivers' behavior (table 11), which choose most frequently to hold constant the average return rate (Figure 17).

Overall, within group preferences seem to be the main driver of “Median” receivers’ behavioral variations.

Advantaged receivers tend to increase the mean return rate with senders of a lower state. Specifically, they significantly increase their return rate when matched with “Disadvantaged” senders (table 11). In this case, they either hold constant (60%) or increase (40%) their return rate (figure 17). A smaller positive effect is observed when “Advantaged” receivers get information on the “Median” state of the counterpart (table 11): i.e. the majority of them (58%) increases the return rate, while the remaining part chooses to either hold constant (30%) or reduce (12%) it (figure 17). Finally, the information on the “Advantaged” state of the sender does not have significant impact over “Advantaged” receivers’ behavior (table 11). This behavior is consistent with the hypothesis of solidarity towards “lower” categories of players.

4.3.6.1 Differences in receivers’ behavior between Trust game 2 and Trust game 3 - Econometric analysis

An ordered probit model is estimated to identify the differential impact of the information about the counterpart’s state on receivers’ behavior between the nine possible matchings, with “Disadvantaged *vs* Disadvantaged” as the default (dropped) variable. The general upshot is that “Disadvantaged” and “Median” receivers are more generous with senders of the same state than with players of different categories. This supports the hypothesis of within-group preferences explaining the behavioral variations of these two categories of players. Furthermore, “Advantaged” receivers are more generous toward “lower” categories of players.

Table 12. Determinants of receivers' behavioral variations between TG2 and TG3

| Difference in the level of transfers back between TG2 & TG3 | Coefficient | Standard error |
|---|-------------|----------------|
| Median <i>vs</i> Disadvantaged | -0.437** | 0.219 |
| Advantaged <i>vs</i> Disadvantaged | -1.028*** | 0.196 |
| Disadvantaged <i>vs</i> Median | -0.399** | 0.186 |
| Median <i>vs</i> Median | 0.639*** | 0.233 |
| Advantaged <i>vs</i> Median | -0.718*** | 0.180 |
| Disadvantaged <i>vs</i> Advantaged | 0.175 | 0.190 |
| Median <i>vs</i> Advantaged | -0.109 | 0.189 |
| Advantaged <i>vs</i> Advantaged | -0.678*** | 0.211 |
| Pseudo R-Squared | 0.048 | |
| n | 720 | |

The negative coefficients associated to the matchings “Median *vs* Disadvantaged” and “Advantaged *vs* Disadvantaged” mean that “Disadvantaged” receivers are more generous with senders of the same state than with “Median” and *particularly* with “Advantaged” senders: if only reciprocal or altruistic motivations drive the decision to give something back to the first player, the information on senders’ “Median” and “Advantaged” state indicates a disparity in the monetary endowment, reducing “Disadvantaged” receivers’ trustworthiness. Indeed, the significant differences between the relevant coefficients indicate that “Disadvantaged” receivers are less trustworthy with “Advantaged” than with “Median” senders.

The behavioral pattern of “Median” receivers mimics the one of “Median” senders. Median receivers exhibit a significantly lower level of trustworthiness towards “Disadvantaged” and “Advantaged” senders compared to the default matching. However, they are less trustworthy with “Advantaged” than with “Disadvantaged” senders. Furthermore, “Median” receivers exhibit the same level of trustworthiness towards group members as in the default matching. This entails that “Median” recei-

vers are significantly and substantially more trustworthy towards senders in the same state than with other categories of players. Hence, solidarity together with within group preferences seem to be the main drivers of “Median” receivers’ behavior.

Advantaged receivers give back significantly less to “Advantaged” senders compared to the default matching. Furthermore, although the coefficients associated to “Disadvantaged *vs* Advantaged” and “Median *vs* Advantaged” do not statistically differ from the default matching, the related difference documents that “Advantaged” receivers are more trustworthy with “Disadvantaged” than with “Median” senders. Hence, “Advantaged” receivers are more generous with the lowest category of players and they modulate their other-regarding behavior on the grounds of the counterpart’s relative position. This result is consistent with the observed positive variation of “Disadvantaged” senders’ offers when they know to be matched with “Advantaged” receivers.

4.4 Conclusions

The chapter reported the results of an experiment aiming at identifying the impact of the random variations of subjects’ state (i.e. monetary endowment) on behavioral trust and trustworthiness in three different economic contexts, where economic heterogeneity is exogenously determined through: 1) positive shocks to total wealth; 2) Redistribution of total wealth; 3) Negative shocks to total wealth. Also, the experiment identified the effect of the information on the counterpart’s state over subjects’ trusting and trustworthy behavior.

In the first trust game, the experiment replicated the evidence of behavioral trust and trustworthiness (Johnson and Mislin 2011) as senders gave approximately 40% of

their endowment on average and receivers reciprocated senders' trusting choice, guaranteeing a return from the transaction.

In the second trust game, heterogeneity determined through redistribution and negative shocks to total wealth reduced the average level of trust, particularly of "Disadvantaged" senders. Specifically, "Disadvantaged" senders are more sensitive (in a negative sense) to economic heterogeneity in a "poor" context rather than in a "richer" one. The effect of heterogeneity on trust is, however, asymmetric across the "Disadvantaged" and "Advantaged" states as "Disadvantaged" senders systematically reduced the level of net transfers, while "Advantaged" ones did not exhibit a univocal pattern of behavior. As it concerns second players, economic heterogeneity disrupted receivers' reciprocity, with a negative effect which is decreasing in receivers' relative position from "Disadvantaged" to "Advantaged". However, the experiment delivered evidence of no significant variations in receivers' behavior across the three relevant economic contexts.

In the third trust game, the information on the "Disadvantaged" or "Advantaged" state of the counterpart induced "Disadvantaged" senders to increase the average level of transfers. This result has been interpreted as evidence of the different motivations that can inspire a trusting choice. Indeed, disadvantaged senders may have increased their offers when matched with disadvantaged receivers because of within-group preferences towards players of the same category. Moreover, the information on the "Advantaged" state of the receiver induced "Disadvantaged" senders to increase the level of transfers to earn the potential profits from the transaction and compensate the loss due to the random penalization in the second trust game. This interpretation found further support in the evidence on receivers' behavior. "Disadvantaged" receivers increased the (average) return rate when they knew to be interacting with a "Disadvantaged" sender, but they reduced it when they were informed on the "Advantaged" state of the sender. Since only reciprocal or altruistic mo-

tivations can inspire receivers' behavior, the information on senders' "Advantaged" state indicated a disparity in the monetary endowment of players, reducing the average level of "Disadvantaged" receivers' trustworthy behavior. Finally, "Advantaged" receivers increased the level of transfers when they knew to be interacting with median and particularly with "Disadvantaged" senders. However, the information on the "Advantaged" state of the sender did not significantly affect their behavior. This behavior is compatible with the hypothesis of solidaristic preferences, inducing "Advantaged" receivers to increase the average level of behavioral trustworthiness when they knew to be matched with "lower" categories of players.

4.5 References

- Anderson, L. R., Mellor, J. M., & Milyo, J. (2006). Induced heterogeneity in trust experiments. *Experimental Economics*, 9(3), 223-235.
- Ashraf, N., Bohnet, I., & Piankov, N. (2006). Decomposing trust and trustworthiness. *Experimental Economics*, 9(3), 193-208.
- Arrow, K. J. (1972). Gifts and exchanges. *Philosophy & Public Affairs*, 343-362.
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1), 122-142.
- Engel, C. (2011). Dictator games: a meta study. *Experimental Economics*, 14(4), 583-610.
- Eurofound. (2012). *Young people not in employment, education or training: Characteristics, costs and policy responses in Europe*. (Dublin, Eurofound).
- Fehr, E., Fischbacher, U., Von Rosenblatt, B., Schupp, J., & Wagner, G. G. (2003). A nationwide laboratory: examining trust and trustworthiness by integrating behavioral experiments into representative survey. *IZA Discussion paper series*, No. 715.

- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817-868.
- Gambetta, D. (2000). Can we trust trust?, in (D. Gambetta, ed.), *Trust: Making and Breaking Cooperative*, Basil Blackwell.
- Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring trust. *Quarterly Journal of Economics*, 811-846.
- Greiner, B., Ockenfels, A., & Werner, P. (2012). The dynamic interplay of inequality and trust — An experimental study. *Journal of Economic Behavior & Organization*, 81(2), 355-365.
- Hardin, R. (1992). The street-level epistemology of trust. *Analyse & Kritik*, 14(2), 152-176.
- Heap, S. P. H., Tan, J. H., & Zizzo, D. J. (2013). Trust, inequality and the market. *Theory and Decision*, 74(3), 311-333.
- Houser, D., Schunk, D., & Winter, J. (2010). Distinguishing trust from risk: An anatomy of the investment game. *Journal of Economic Behavior & Organization*, 74(1), 72-81.
- Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5), 865-889.
- Putnam, Robert; Robert Leonardi, Raffaella Nanetti (1993). *Making Democracy Work*. (NJ, Princeton University Press).
- Sapienza, P., Toldra-Simats, A., & Zingales, L. (2013). Understanding trust. *The Economic Journal*, 123(573), 1313-1332.
- Xiao, E., & Bicchieri, C. (2010). When equality trumps reciprocity. *Journal of Economic Psychology*, 31(3), 456-470.

General Conclusions

This thesis studied experimentally the extent to which trusting and trustworthy behavior depend on individuals' state (e.g. social, economic status) as well as the way the knowledge of the counterpart's state mediates behavioral trust and trustworthiness. To pursue this aim, the first experiment implemented two trust games with young people in different labor market states (i.e. employed, student and NEET) and fed subjects with information on the counterpart's employment state in the second game. To allow for causal inferences, the second experiment implemented three trust games. In the first one, subjects' play the game in condition of equality (i.e. all players got the same monetary endowment), while in the second one a random mechanism determined a distribution of "Disadvantaged", "Median" and "Advantaged" players through: Positive shocks to total wealth (Treatment 1); Redistribution of total wealth (Treatment 2); Negative shocks to total wealth (Treatment 3). In the third trust game, subjects' relative position in the distribution was the same as that one determined in the second game, but they were informed of the counterpart's state.

As the first Chapter clarified, individuals' motivations (i.e. preferences and beliefs) to trusting and trustworthy behavior are likely to vary with their state and with the information on the counterpart's state. Since no theoretical priors provide an univocal prediction of the effect of the random change in the economic state on behavioral trust and trustworthiness, the proposed experiments aimed at identifying behavioral patterns, and the related causes, which still need a theoretical explanation.

The first experiment showed that unemployed people (i.e. NEETs searching for a job *without* success) were the least trustworthy category. This results is consistent with the negative effect of the information on the NEETs status of the receiver on employed senders' trust. These findings jointly suggest a candidate mechanism for the micro-level explanation of the persistence of long-term wage and employment pe-

nalties: The negative effect of the unemployment state over young individuals' reliability provides a rationale for employed people to distrust them; this, in turn, may impede a successful search for a job in the long run to young unemployed people. Moreover, subjects in precarious forms of employment exhibited systematically lower levels of trust/trustworthiness. The implication is that the increasing tendency to promote flexible contracts is likely to have long-term negative consequences for young people's labour market, operating through the channel of the negative effects of precarious employment on young people's trust/trustworthiness. Finally, NEETs tended to reduce their average level of trust/trustworthiness when they got information on the employed and student status of the counterpart, while they significantly increased it when they knew to be interacting with other NEETs. This is consistent with the hypothesis of within-group preferences inducing disadvantaged subjects to a more cooperative behavior towards group members at the expenses of outliers.

The second experiment documented an asymmetry between the effects of the "Disadvantaged" and "Advantaged" states, with the former strongly reducing subjects' behavioral trust and trustworthiness, and the latter having an ambiguous and noisy impact on behavior. Aggregating across all states, heterogeneity induced through redistribution and negative shocks to total wealth significantly reduced the average level of trust, while positive shocks to total wealth did not have any aggregate effect on behavior. This difference was mainly driven by the behavioral variations of "Disadvantaged" senders, which were more sensitive (in a negative sense) to heterogeneity in "poor" economic contexts (i.e. Treatment 3) than in "richer" ones (i.e. Treatment 1). Similarly, induced heterogeneity reduced the average level of trustworthiness of all the categories of receivers, particularly of "Disadvantaged" ones.

In the third trust game, the knowledge of players' state induced some behavioral inconsistency with respect to the second one. Specifically, "Disadvantaged" senders increased their level of trust when they knew to be playing with "Disadvantaged"

and “Advantaged” receivers, though they were more generous with the second category. This supports the characterization of trusting and trustworthy behavior in terms of multiple and non-exclusive motivations. On the one hand, “Disadvantaged” senders increased the average level of trust when matched with “Disadvantaged” receivers because of altruism towards individuals in the same category. On the other hand, the “Advantaged” state of the receiver reduced the cost of reciprocity, so to provide a rationale for “Disadvantaged” senders to expect higher profits from the transaction and consequently to increase their offers. The results on “Disadvantaged” receivers’ supported this interpretation. “Disadvantaged” receivers increased the average level of transfers back when they knew be matched with “Disadvantaged” senders because of within group preferences, but they reduced it in the case of matching with “Advantaged” senders. Since only (conditional or unconditional) other-regarding preferences motivate a positive level of net transfers from receivers, the information on senders’ “Advantaged” state indicated a disparity in players’ endowment, inhibiting their reciprocal or altruistic behavior. Information on the counterpart’s state did not significantly impact over “Advantaged” senders’ behavior, while “Advantaged” receivers exhibited a solidaristic behavior as they increased their level of trustworthiness only when they played the game with lower categories of senders (i.e. “Median” and “Disadvantaged”).

Both experiments provided similar results on the behavioral variations of the most disadvantaged players: i.e. unemployed NEETs in experiment 1 and “Disadvantaged” in experiment 2. Indeed, in both cases players’ heterogeneity had a stronger negative effect on disadvantaged players. This support the hypothesis of a causal relationship between the unemployment state and the reduced level of trust. Moreover, both NEETs and “Disadvantaged” players increased the level of trust and trustworthiness when they got information on a counterpart in the same state. This result documented the robustness of within group preferences with respect to the way the

disadvantaged state is determined: as the output of the labor market in the first experiment and as an exogenously induced state in the second one.

The evidence accruing from the two experiments suggests three main avenues of future research. The design of the first experiment can be easily integrated with field data to examine the the cross-country differences in behavior and the interaction between country context and behavioral differences across labor market states. This may be further supported by experiments in different locations within the same countries so as to control for country fixed effects arising out of culturally determined social norms.

The results and the related interpretation of the second experiment require an analysis of both senders' and receivers' beliefs to better identifying the idiosyncratic contribution of the multiple and non-exclusive motivations to behavioral trust and trustworthiness. Moreover, data on players' beliefs together with an index of self-perceived trustworthiness would be the empirical basis to test the introspection hypothesis, according to which senders expect receivers to behave as they would if they were to play that role; these introspection-based beliefs are likely to drive their trusting behavior.

Finally, to increase the realism of the experimental setting and reducing the differences between labor market and exogenously induced states, the design of the second experiment could be integrated with a real effort task determining the distributions of players' categories ("Disadvantaged", "Median", "Advantaged"). This would allow to assess the way the subject-specific level of effort mediates the effect of players' relative position as well as the effect of the information on the counterpart's state on behavioral trust and trustworthiness.