EconomEtica

Centro interuniversitario per l'etica economica
e la responsabilità sociale di impresa
promosso dalla Fondazione Italiana Accenture

N.2 July 2008

Marco Faillo,
Stefania Ottone e
Lorenzo Sacconi

Compliance by believing: an experimental exploration on social norms and impartial agreements

Working papers

Compliance by believing: an experimental exploration of social norms and impartial agreements.

Marco Faillo
(University of Trento)

Stefania Ottone
(University of Milano-Bicocca and EconomEtica)

Lorenzo Sacconi
(University of Trento and EconomEtica)

Abstract

The main contribution of this paper is twofold. Firstly, it focuses on the decisional process that leads to the creation of a social norm. Secondly, it analyses the mechanisms through which subjects conform their behaviour to the norm. In particular, our aim is to study the role and the nature of Normative and Empirical Expectations and their influence on people's decisions. The tool is the Exclusion Game, a sort of 'triple mini-dictator game'. This is a situation where 3 subjects – players A - have to decide how to allocate a sum S among themselves and a fourth subject - player B - who has no decisional power. The experiment consists of three treatments. In the Baseline Treatment participants are randomly distributed in groups of four players and play the Exclusion Game. In the Agreement Treatment participants in each group are invited to vote for a specific non-binding allocation rule before playing the Exclusion Game. In the Outsider Treatment, following the voting procedure and before the Exclusion Game is played, a player A for each group (the outsider) is reassigned to a different group and instructed on the rule chosen by the new group. In all the treatments, at the end of the game and before players are informed about the decisions taken during the Exclusion Game by the other co-players, first order and second order expectations (both normative and empirical) are elicited by means of a brief questionnaire. The first result obtained is that subjects' choices are in line with their empirical (not normative) expectations. The second result is that even a non-binding agreement induces a convergence of empirical expectations – and, consequently, of choices. The third result is that expectation of conformity is higher in the partner protocol. This implies that a single outsider disrupts the 'conditional compliance' equilibrium.

Keywords: fairness, social norms, beliefs, psychological games, experimental games.

JEL Classification: C72, C91, A13.

Acknowledgements: we thank the staffs of the Computable and Experimental Economics Laboratory of the University of Trento (CEEL) and the Experimental Economics Laboratory of the University of Milano-Bicocca (EELAB) for their technical support. Financial support from EconomEtica and the Department of Economics of the University of Trento is gratefully acknowledged. We thank Martin Dufwenberg, Tore Ellingsen, Michele Bernasconi, Enrique Leonidas de la Rosa for their useful comments. We also thank Laura Magazzini and Cinzia Di Novi for their useful suggestions concerning the econometric analysis. The usual disclaimers apply.

This paper has been presented at the SIDE conference in Milan – Bocconi (September 2007), at the IMEBE 2008 conference in Alicante (March 2008) at the Urrutia Elejalde XI Summer School on "Economics and Philosophy: Social Norms" in San Sebastian (July 2008), at the IAREP/SABE conference in Rome (September, 2008), at the ESA conference in Lyon (September 2008) at the International Summer School on "Why isn't Ethics a Behavioral Science ?" in Trento (September 2008) and at the 3rd Nordic Conference on Behavioral and Experimental Economics in Copenhagen (November 2008).

Introduction

In recent years the concept of social norm has become very popular among behavioral and experimental economists interested in deviations from purely selfish behavior. Notwithstanding the increasing attention paid to the concept, a systematic study of social norms as motivational drivers for economic agents is still lacking. At empirical level, no or little effort has been made to find a way to observe the creation of social norms in the laboratory. At theoretical level, the social preferences (Bolton, 1991; Bolton and Ockenfelds 1998; Fehr and Schmidt, 1999; Charness and Rabin, 2002) and reciprocity models (Rabin, 1993; Falk and Fischbacher, 2006) incorporate normative principles within individual preferences instead of treating them as separate objects. This prevents explanation of compliance with a social norm in a context where the norm prescribes a choice which implies a cost in terms of material self interest, where external enforcement (through sanctions or rewards) is not feasible (for example, due to imperfect monitoring), and where reputational mechanisms cannot be effective due to ex-post non-verifiability or to the fact that the interaction is one-shot (Faillo and Sacconi, 2007).

In what follows we seek to fill these gaps by describing an experimental procedure in which a norm is created by means of an ex-ante impartial agreement among the members of a group. In particular, we consider a context in which the norm arises from agreement on a rule on how to play a one-shot game which follows the agreement, and in which subjects must decide how to divide a sum of money among themselves. The agreement is reached by subjects who vote for the rule behind a veil of ignorance, i.e. before they know what their role in the game will be. By means of this procedure we reproduced in the laboratory a situation in which the subjects first participated in a "constitutional" stage in definition of a non-binding norm and then discovered their roles in the game and decided whether or not to comply with the norm.

We explain the subjects' decision to comply with the norm in this specific context in two steps. First, we posit a "conditional compliance" hypothesis, according to which, in a strategic interaction among N players, player *i's*

decision to comply with a shared norm dictating a choice in contrast with her material self-interest depends on her beliefs about the other *N-1* players' choices and beliefs.

Second, assuming that the "conditional compliance" hypothesis holds, we examine the agreement's role in inducing compliance by the subjects. In this regard, we can distinguish between two approaches to the question of how the agreement – which in standard non cooperative game theory is just 'cheap talk' – influences the subjects' decision to comply. According to the first approach, the agreement performs a key role, and if it is associated with the appropriate expectations of reciprocal compliance, it induces subjects to comply. According to the second approach, the agreement is not important in itself. What matters is the presence of expectations about the existence of a norm (which may also originate from experience, education, communication etc.) which constitute the basis for the emergence of normative expectations .

A further question to be addressed concerns the relationship between the agreement and the emergence of expectations of reciprocal conformity. Sacconi and Faillo (2008) show that, for a significant number of subjects, the agreement is a sufficient condition to expect reciprocal conformity, and therefore to comply with the rule. As discussed elsewhere (Faillo and Sacconi, 2007), an appropriate explanation for this finding cannot be based on a standard logical inference – according to which the existence of expectations of reciprocal conformity derives from the existence of the agreement – but should be based on non-monotonic logic and default reasoning.

The foregoing discussion can be further developed by looking at the few attempts made in the behavioral economics literature to devise a theory on conformity with social norms. In particular, we can identify two alternative interpretations of the "conditional compliance" hypothesis, which differ also in terms of the role assigned to the impartial agreement in explanation of norm compliance. Sacconi and Grimalda (2007) develop a model of conformist preferences based on psychological game theory. According to this model, a player characterized by conformist preferences complies if she

participates in choosing the norm in a social contract setting (i.e. she participates in an ex-ante agreement on the norm), if she expects that other players who have contributed to choose the rule will comply (First Order Empirical Expectations) and if she expects that others will expect that she will comply (Second Order Empirical Expectations).

Cristina Bicchieri (2006) argues that compliance is observed when the player is aware of the norm's existence (the agreement is not necessary) and believes that a sufficiently large number of people comply with the norm (First Order Empirical Expectations); *and* either a sufficiently large number of people think that she ought to conform *or* a sufficiently large number of people are ready to sanction her for not conforming (Second Order Normative Expectations).

Bicchieri and Xiao (2007) run an experiment in which they show that, when normative expectations (what we believe others think ought to be done) and empirical expectations (what we expect others will actually do) are in contrast, subjects choose according to the latter.[1]

There are apparent similarities, but also important differences, between these two theories and the guilt aversion model (Charness and Dufwenberg, 2006; Battigalli and Dufwenberg, 2007). According to the guilt aversion hypothesis, people care about what others expect them to do and feel guilty if they do not fulfill what they think are others' expectations. As in conformist preferences and Bicchieri's theory, people's behavior depends on their second order (empirical) beliefs; but differently from these two theories, in the guilt aversion model first order beliefs are not considered as drivers of individual choice.

In order to take account of the alternative interpretations of the "conditional compliance" hypothesis, in our experiment we investigated four types of expectations of a generic player *i*:

[1] Further evidence on the role of empirical and normative expectations in fostering compliance with norms of fairness can be found in a recent paper by Krupka and Weber (2007).

First Order Empirical Expectations (FOEE): player *i*'s beliefs about the other players' choices.
Second Order Empirical Expectations (SOEE): player *i*'s beliefs about other players' beliefs about her choice.
First Order Normative Expectations (FONE): player *i*'s beliefs about what is the right choice in a particular situation.
Second Order Normative Expectations (SONE): player *i*'s beliefs about what other players consider as the right choice in a particular situation.

We studied the effect of the agreement on the emergence of different types of expectations, and the contribution of these different types of expectations to explanation of the decision to comply with a shared norm. We considered a simple game, and we started by studying the relationship between choice and expectations. To this end, we observed how the subjects played the game and we collected data on what they believed others would do and expect. We then added analysis of how the introduction (before the actual playing of the game) of an agreement on a non-binding division rule influenced the subjects' expectations, and consequently the way in which the game was played. Finally, we considered the case in which subjects played the game with co-players who were not those with whom they had participated in the agreement.
As will become clearer below, these steps corresponded to the three treatments of our experimental design: the *Baseline Treatment* (BT), the *Agreement Treatment* (AT), and the *Outsider Treatment* (OT). The BT gave us general information about the relationship between choice and empirical and normative expectations. The comparison between what we observed in BT and AT enabled us to examine the influence of the agreement on expectations and choice and to identify the roles of the different types of expectations in inducing compliance with the norm. Finally, by comparing the AT with OT we could assess whether being in a group with individuals who had not been their partner in the agreement influenced the subjects' decision to comply.

The paper is organized as follows. Section 2 describes the experimental design, procedure and hypotheses, while Section 3 analyses the results. A discussion of the results and some concluding remarks end the paper (Section 4).

2. Experimental Design

The tool used was the *Exclusion Game* (Sacconi and Faillo, 2005; Faillo and Sacconi, 2007), a sort of 'triple mini-dictator game'. This is a situation where 3 subjects – players A (A1, A2 and A3 respectively) – must decide how to allocate a sum S among themselves and a fourth subject – player B – who has no decisional power. In particular, A1, A2 and A3 have to decide separately and independently the amount that they want to request for themselves, choosing one of three possible strategies: asking for 25%, 30% or 33% of S. The payoff for players A is exactly the sum requested for themselves (a1, a2 and a3 respectively), while the payoff for player B is the remaining sum (S – a1 – a2 – a3). In our experiment, each group was given 60 tokens – each token corresponded to € 0.50 – and each player A's strategies were : "Ask for 15 tokens", "Ask for 18 tokens", "Ask for 20 tokens".

The experiment consisted of three treatments: the *Baseline Treatment* (*BT*), the *Agreement Treatment (AT)* and the *Outsider Treatment (OT)*.

In the *Baseline Treatment* participants were randomly distributed in groups of four players and played the *Exclusion Game*.

In the *Agreement Treatment* participants were randomly distributed in groups of four players and were told about the stages of the experiment and about the *Exclusion Game*. In the first stage, without knowing their role in the game, they took part in a voting procedure. In each group participants were invited to vote for a specific allocation rule. In particular, subjects had to vote for one of three alternative rules (the fourth number was the type B player's payoff): {15,15, 15,15},{18,18, 18,6}, {20,20, 20,0}. The first rule assigned the same payoff to every member of the group; the second rule corresponded to a partial inclusion of player B in the share-out the money; the third rule implied the

total exclusion of the type B player.[2] Players had to reach a unanimous agreement on the rule within a limited number of trials (10 in our experiment). Voting was computerized and completely anonymous. The agreement was not binding, but failure in reaching it was costly, since only groups who reached agreement in this first stage could participate in the second stage. In the second stage the composition of the groups was unchanged and roles were randomly assigned to implement the *Exclusion Game*. In this case, players A could either decide to implement the rule selected by vote or choose one of the alternative allocations. Players who did not enter the second stage waited until the end of the session. Their payoff was the show-up fee.

In the *OT* participants were randomly distributed in groups of four players, and they were instructed on the stages of the experiment and on the *Exclusion Game*. The first stage, as well as the rule on entering the second stage, were the same as in the *AT*. At the beginning of the second stage, players were informed about their role, and groups were rematched. In particular, a player A for each group (the outsider) was reassigned to a different group and told about the rule chosen by the new group, while the other members of the group did not know what rule the outsider's previous group had adopted. After the re-matching, the subjects played the *Exclusion Game*. For a summary of the three treatments see Figure 1.

[2] Note the correspondence between the rules and the strategies of the Exclusion Game: for players A compliance with the {15,15,15,15} rule implies choice of the "ask for 15 tokens" strategy; compliance with the {18,18,18,7} rule implies choice of the "ask for 18 tokens" strategy; and, finally, compliance with the {20,20,20,0} rule implies choice of the "ask for 120 tokens" strategy

2.1 Experimental Procedure.

The experiment was run in both Milan (EELAB – University of Milan Bicocca) and Trento (CEEL – University of Trento) [3]. We ran 3 sessions for the *BT* (1 in Milan and 2 in Trento), 4 sessions for the *AT* (2 in Milan and 2 in Trento), 5 sessions for the *OT* (3 in Milan and 2 in Trento). Overall, 216 undergraduate students – 104 in Milan and 112 in Trento – participated in the experiment. A more detailed description of the sessions is given in Table 1.

The experiment was programmed and conducted using the z-Tree software (Fischbacher, 2007). The instructions were read by participants on their computer screen while an experimenter read them out loud.[4]

After the instructions had been read, and before the subjects were invited to take their decisions, some control questions were asked in order to ensure that the players had understood the rules of the game. At the end of each session, subjects were asked to fill in a questionnaire for the collection of socio-demographic data.

Players were given a show-up fee of 3 euros.

2.2 Beliefs elicitation.

In all the treatments, at the end of the game and before players were informed about the decisions taken during the *Exclusion Game* by the other co-players, first order and second order expectations (both normative and empirical) were elicited by means of a brief questionnaire. In particular, in each group each player made statements concerning:

1. the probabilities of each possible choice by co-players A (First Order Empirical Expectations);

[3] At the University of Trento, subjects were recruited by posting ads in various departments. Ads were posted one week before the experiment. Enrolments by students interested in taking part in the experiment were collected by the staff of the Computable and Experimental Economics Laboratory (CEEL) of the University of Trento.
At the University of Milano-Bicocca subjects were recruited by email. They were students on the mailing list of the Experimental Economics Laboratory of the University of Milano-Bicocca (EELAB). Two weeks before the experiment they received an email inviting them to visit the Laboratory's website for information about the experiment and subscriptions.

[4] The instructions and the program are available upon request.

2. the probability of each co-player's possible judgement about her own choice (Second Order Empirical Expectations);
3. the choice she considered to be the 'right' one A (First Order Normative Expectations);
4. the choice that co-players considered to be the 'right' one (Second Order Normative Expectations).[5]

In the *OT*, guesses about the behaviour and beliefs of partners and outsiders were elicited separately.

Only good guesses of the Empirical Expectations were rewarded on the basis of a quadratic scoring rule (Davis and Holt, 1993).[6]

2.3 Experimental Hypotheses.

Hypothesis 1 (H1): According to psychological game theory,[7] people's preferences depend on their expectations (of different orders and nature). Consequently, the players' choices in the *Exclusion Game* could be explained in terms of their expectations about the behaviour of the other players.

Moreover, if Bicchieri and Xiao (2007) are right, when normative and empirical expectations are in contrast, the latter play a more important role in players' decision-making.

Hypothesis 2 (H2): In treatments AT and OT, agreement should be reached by all the groups because it is not binding but its failure is costly (failure would prevent the players from entering the second stage of the experiment).

Hypothesis 3 (H3): The possibility of agreeing with a distributive norm enhances compliance by inducing a convergence of individual expectations. In

[5] See appendix 1 for details on the belief elicitation procedure.

[6] We used the following scoring rule:

$$Q(p) = a - b\sum_{k=1}^{N}(I_k - p_k)^2$$

where I_k takes value 1 if the event realized is event k and 0 otherwise. p_k is the probability associated with event k. The maximum score is a, and the minimum score is a-$2b$. We chose $a=2$ e $b=1$.

[7] See for example Geanakoplos et al. (1989); Rabin, (1993); Dufwenberg (2006)

other words, compliance can be explained in terms of the emergence of reciprocal expectations of conformity due to the agreement (this hypothesis is compatible with both Sacconi and Grimalda's theory of conformist preferences and Bicchieri's theory of social norms).

Hypothesis 3a (H3a):, subjects will comply if i) they believe that other members of their group will comply (First Order Empirical Expectations compatible with the choice dictated by the rule) *and* if ii) they believe that other members of the group think that complying is the right thing to do (Second Order Normative Expectations compatible with the choice dictated by the rule) (this hypothesis is compatible with Bicchieri's theory)

Hypothesis 3b (H3b): subjects will comply if i) they participate in the agreement on the rule; ii) they believe that other members of their group will comply (First Order Empirical Expectations compatible with the choice dictated by the rule) *and* if iii) they believe that other members of the group expect that they will comply (Second Order Empirical Expectations compatible with the choice dictated by the rule). With respect to point (i), compliance should be less frequent in the OT treatment, where groups are re-matched (this hypothesis is compatible with Sacconi and Grimalda's theory of conformist preferences).

3. Data analysis

In this section we provide an overview of our experimental data and results by discussing two main points. First, we analyse the relation between beliefs and behaviour: in particular, we shall check whether beliefs influence the subjects' decision-making process. Second, we test whether and how different scenarios influence beliefs and, consequently, people's decisions.

3.1 Description

Overall, 216 undergraduate students took part in the experiment. 56 players were recruited for the BT, 72 for the AT, and 88 for the OT. We have observations of 42 subjects A in the BT, 54 in the AT, and 66 in the OT.

In the BT, the majority of players A chose to ask for the highest amount of tokens (20) – 73.8% against the 21.4% who chose 18 and the 4.8% who chose 15. The situation is different in both the AT and the OT. In the AT, 37%, 16.7% and 46.3% chose respectively 20, 18 and 15; inn the OT the percentages are 54.5%, 12.1% and 33.4%.

As regards the rule voted for, the 15-15-15-15 rule seems to have been the preferred option in both the AT and the OT. In particular, 17 groups out of 18 in the AT and 20 out of 22 in the OT chose the fair-division rule. The 18-18-18-6 rule was chosen by 1 group in the AT and by 1 group in the OT; only 1 group in the OT chose the 20-20-20-0 rule. 50% of players in the AT and 39.4% in the OT complied with rule selected when playing the *Exclusion Game*.

3.2 Results

Result 1. The subjects' choices are in line with their expectations.

On checking for correlation between beliefs and decisions, we find that most of the players' choices are in line with either empirical or normative expectations (Table 2).[8] However – as in Bicchieri and Xiao (2007) – when normative and empirical expectations are in contrast, the latter play a more important role in the players' decision-making (Table 3) and they are significantly correlated to the subjects' choices (Spearman test; $p < 0.03$). This is not to be the case when we analyse normative expectations (Spearman test; $p > 0.17$).[9]

[8] We consider only first order expectations since second order expectations are either equal or highly correlated to the former. For a more detailed description, see Appendix 1.

[9] Test run only on observations where FONE and FOEE are different.

Result 2. When agreement is possible, it reached by all groups. Almost all groups agree on the 15-15-15-15 rule.

As we expected, when agreement is possible, it is reached by all groups. This is a quite obvious result: agreement is not binding but failure in reaching it is costly. However, the interesting point is that the fair rule 15-15-15-15 seems to be a sort of focal point (see Table 4). What this means can be explained by looking at the results of the first voting attempt. Table 5 shows that the first choice of 75% of players in the AT and 70% of players in the OT is the 15-15-15-15 rule. On running a binomial test (choosing the 15-15-15-15 rule against choosing another rule) we find that these values are significant (p = 0.000 in the AT and p = 0.04 in the OT). This may imply that most subjects knew perfectly well what was right. However, what happened to the remaining 25% and 30%? Why did most of them changed their minds? And why, when playing the *Exclusion Game,* did 50% of subjects in the AT and 61% in the OT decide not to comply with the rule (Table 6)? A possible explanation is that 'unfair' subjects voted for the non-binding 'fair' rule in order to end the time-consuming voting procedure. However, this was not enough for players who preferred the 'fair' rule. They knew perfectly well that the agreement was not binding (in fact, among players who eventually voted for a rule different from their first choice, 71% did not comply with it when playing the *Exclusion Game*) and if they thought that the other co-players were not complying, they probably defected as well. This would be in line both with the fact that empirical expectations are more important than normative ones, and with the higher probability of expecting the others to choose 20 (at least in the AT) as soon as the number of voting rounds increased (see Appendix 2).

Result 3. Agreement induces convergence of empirical expectations.

In the BT at least 70% of the players asked for 20, while in the AT only 37% of the participants asked for the maximum. This difference is significant

(Mann-Whitney[10]; $p = 0.0002$). However, our experimental hypothesis is more complicated and implies that our participants engage in a two-step reasoning process. Step 1: the agreement influences the players' empirical expectations. Step 2: empirical expectations define the subjects' choices. This means that we want to show that the difference between BT and AT is a consequence of the impact of the agreement on players' beliefs and preferences.

In the AT, 17 groups out of 18 chose the 15-15-15-15 rule, and 1 chose the 18-18-18-6 one. On analysing the subjects' expectations, we find that in the AT there is a significant decrease of subjects who think that the other members of their group have asked for 20 tokens (Table 7). A probit regression – where the dependent variable is the probability of expecting the others to choose 20 – shows that subjects are more likely to expect selfish behaviour by the co-players in the BT ($p = 0.000$). A bivariate recursive probit confirms both the influence of beliefs on the subjects' decisions ($p = 0.00$) and the convergence of empirical expectations on a choice in line with the fair rule ($p = 0.000$).[11] More details on the econometric analysis are given in Appendix 2.

Result 4. Expectation of conformity is higher in the partner protocol.

When we introduce a mixed protocol whereby the *Exclusion Game* is played in groups where one subject is an 'outsider' (in the OT), a lower percentage of players comply with the chosen rule (Table 6). Again, our experimental hypothesis is that our participants engage in a two-step reasoning process. Step 1: the introduction of an outsider influences the players' empirical expectations. Step 2: empirical expectations define the subjects' choices. This means that, once again, we want to show that the difference between AT and OT is a consequence of the impact of the outsider on players' beliefs. If we analyse expectations, it turns out that in the AT players believe in their co-players' compliance more than in the OT (Table 8). A probit regression –

[10] Independent observations are average choices of each group so as to take account of the fact that choices within the same group in the AT are not independent.

[11] This result is perfectly in line with the result obtained by Sacconi and Faillo (2005) using a within-subject design.

where the dependent variable is the probability of expecting the others to comply – shows that subjects are more likely to expect compliance in the AT ($p = 0.046$). A bivariate recursive probit confirms both the influence of beliefs on the subjects' decisions ($p = 0.012$) and the fact that in the OT subjects are more likely to expect deviation by the co-players from the chosen rule. ($p = 0.051$). More details on the econometric analysis are given in Appendix 2.

Result 5. The existence of normative expectations in line with the shared norm is not a necessary condition for compliance.

The previous analyses confirm the robustness of Sacconi and Grimalda's theory. According to hypothesis H3a, First Order Empirical Expectations (FOEE) and Second Order Empirical Expectations (SOEE) should be compatible with the choice dictated by the rule. In our data, SOEE are in line with FOEE (see result 1). Moreover, FOEE influence the subjects' decisions (see result 3 and result 4), and participation in the agreement has a significant impact on the decision to comply (result 4.). This results gives support to Sacconi and Grimalda's theory of conformist preferences.

On the other hand, Bicchieri's theory seems to be less robust. According to hypothesis H3b, both First Order Empirical Expectations (FOEE) and Second Order Normative Expectations (SONE) in line with the rule selected are necessary to predict compliance. To check this point we isolate the subgroup of subjects who comply with the rule selected and whose FOEE are in line with it. We obtain a subgroup of 14 subjects in the AT and 14 subjects in the OT. When we analyse the correlation between SONE and choice it turns out that they are correlated neither in the AT (Spearman correlation coefficient; $p = 0.23$) nor among the insiders in the OT (Spearman correlation coefficient; $p = 0.5$). They are only slightly correlated among the outsiders in the OT (Spearman correlation coefficient; $p = 0.07$), but in this case we have only 6 observations.

4. Conclusions

The aim of this paper has been twofold. Firstly, it has focused on the decision-making process that leads to the creation of a social norm. Secondly, it has analysed the mechanisms through which subjects conform their behaviour to the norm.

We can summarize our results as follows.

1) The subjects' choices are in line with their empirical expectations, and when normative and empirical expectations are in contrast, the latter play a more important role in the players' decisions (H1).
2) Agreement is reached in all groups (H2).
3) Even a non-binding agreement induces convergence of empirical expectations and, consequently, of choices (H3). Moreover, this finding confirms the robustness of the results obtained by Faillo and Sacconi (2008). In particular, it is perfectly in line with the "conditional compliance" hypothesis, according to which subjects comply with a norm if they believe that other members of their group will comply and if they believe that other members of their group expect that they will comply (H3b).
4) The results of the Outsider Treatment seems to suggest that participation in the agreement is a necessary condition for compliance. When groups are rematched and one of the players A (the outsider) is assigned to a new group, the members of her new group (the insiders) do not expect compliance from her, and consequently they do not comply (H3b). The outsider seems to acknowledge this, and, on expecting non-compliance by the insiders, she does not comply.
5) Our last result (a generally non significant correlation between Second Order Normative Expectations and choice of conformity) does not confirm the hypothesis that both first order empirical expectations and second order normative expectations are necessary conditions for compliance (H3a).

References

Battigalli, P. and Dufwenberg, M. (2007). Guilt in games. *American Economic Review, Papers and Proceedings* 97, 170-176.

Bicchieri, C. (2006), *The Grammar of Society*: *The Nature and Dynamics of Social Norms*, Forthcoming Cambridge University Press.

Bicchieri, C., Xiao E., (2007), *Do the Right Thing: But only if others do so*, MPRA Paper 4609, University Library of Munich, Germany.

Bolton, G.E. (1991), "A Comparative Model of Bargaining: Theory and Evidence" *American Economic Revie*w, 81, 1096-1136.

Bolton, G. E, and A. Ockenfels (1998), "Strategy and Equity: an ERC-analysis of the Güth-van Damme game", *Journal of Mathematical Psychology*, 42(2), 215-226.

Camerer, C.F. (2003), *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton; NJ; Princeton University Press.

Charness, G. and Dufwenberg, M. (2006). Promises and partnership. *Econometrica* 74, 1579-1601.

Charness, G. and M. Rabin (2002), "Understanding Social Preferences with Simple Tests", *Quarterly Journal of Economics* , 117(13), 817-869.

Davis, D. D. and Holt, C.A. (1993) *Experimental Economics*. New Jersey: Princeton University Press.

Dufwenberg, M. (2008), "Psychological Games" in Durlauf, S.N. and Blume, L. E. (eds.), *The New Palgrave Dictionary of Economics* (2nd Edition), Palgrave Macmillan.

Di Novi C. (2007), An Economic Evaluation of Life-Style and Air-pollution-related Damages: Results from the BRFSS, JEPS Working Papers 07-001, JEPS.

Faillo, M. and Sacconi L. (2007), "Norm Compliance: The Contribution of Behavioral Economics Theories", in Innocenti, A. and Sbriglia, P. (eds.), *Games, Rationality and Behaviour*, *Essays in Behavioural Game Theory and Experiments,* New York, Palgrave-MacMillan.

Falk, A. and U. Fischbacher (2006), " Theory of Reciprocity" Games and Economic Behavior, 54 (2): 293-315.

Fehr, E. and K.M. Schmidt (1999). "A Theory of Fairness, Competition and Co-operation." Quarterly Journal of Economics, 114: 817-868.

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

Geanakoplos, J., D. Pearce and E. Stacchetti (1989), "Psychological Games and Sequential Rationality", *Games and Economic Behavior*, Vol. 1, pp. 60-79.

Grimalda, G., L. Sacconi (2005) "The Constitution of the Not-for-Profit Organisation: Reciprocal Conformity to Morality", *Constitutional Political Economy,* Vol 16(3), 249-276.

Krupka, E. and Weber, R.A. (2007), **"**The Focusing and Informaitonal Effects of Norms on Pro-Social Behavior", *Discussion Paper No. 3169.*

Rabin M., 1993, 'Incorporating Fairness into Game Theory and Economics',
American Economic Review, 83(5), 1281-1302.

Sacconi, L. and Faillo, M. (2008) "Conformity, reciprocity and the sense of justice. How social contract-based preferences and beliefs explain norm compliance: the experimental evidence", *Discussion Paper no. 14/08, Department of Economics University of Trento*

Sacconi L. and G. Grimalda (2007) "Ideals, conformism and reciprocity: A model of Individual Choice with Conformist Motivations, and an Application to the Not-for-Profit Case" in (L.Bruni and P.L.Porta eds.) *Handbook of Happiness in Economics*, Edward Elgar, London (in print).

Figure 1. Treatments

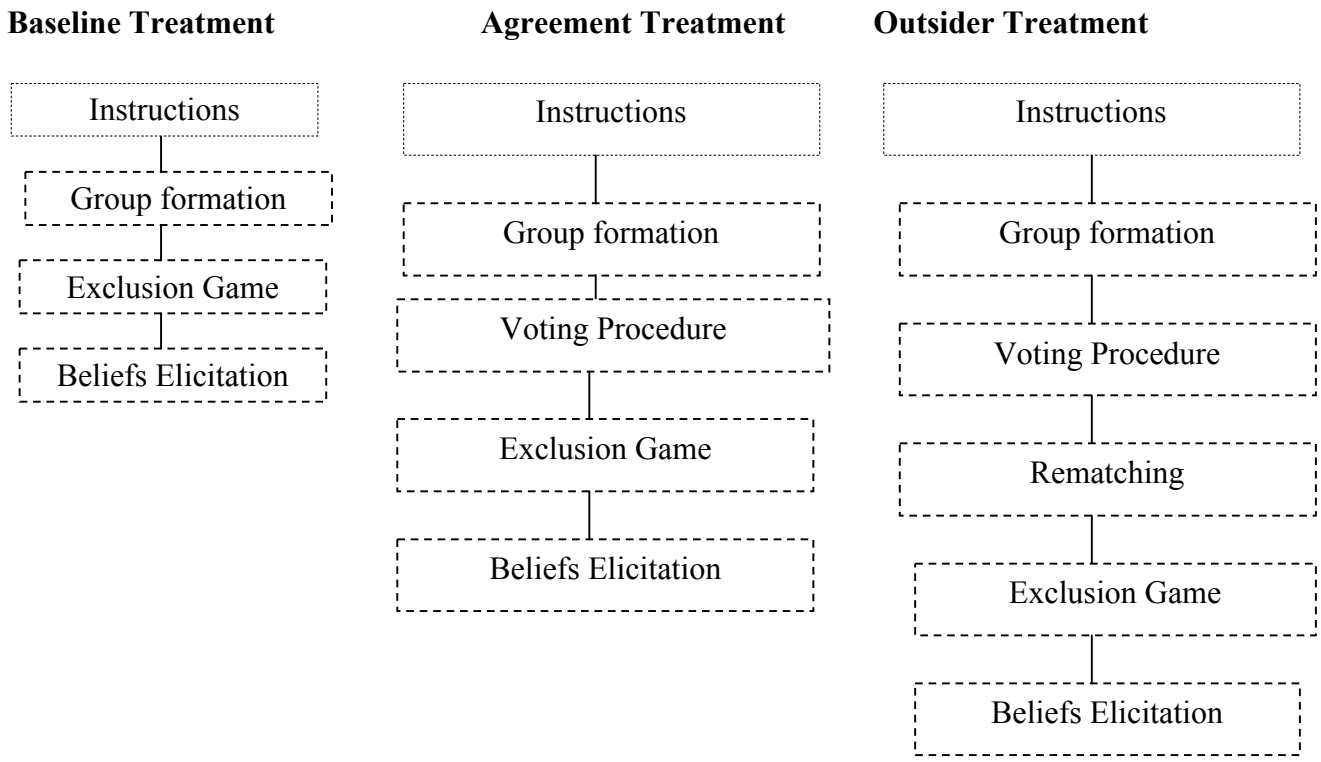

Table 1. Experimental Design

Treatment	Voting Procedure	Matching	Sessions	Subjects
BT	NO	Partner Protocol	2 in Trento (T) 1 in Milan (M)	36 (T) + 20 (M) 9 groups (T) + 5 groups (M) (27 (T) + 15 (M) players A)
AT	YES	Partner Protocol	2 in Trento (T) 2 in Milan (M)	36 (T) + 36 (M) 9 groups (T) + 9 groups (M) (27 (T) + 27 (M) players A)
OT	YES	Mixed – Partner and Stranger Protocol	2 in Trento (T) 3 in Milan (M)	32 (T) + 56 (M) 8 groups (T) + 14 groups (M) (24 (T) + 42 (M) players A)

Table 2. Beliefs and choices

	It is possible to explain subjects' behaviour through…		
	FOEE	FONE	OTHER
BT			
T (N = 27)	82%	7%	11%
M (N = 15)	93%	0%	7%
AT			
T (N = 27)	82%	11%	7%
M (N = 27)	82%	7%	11%
OT			
T (N = 24)	71%	21%	8%
M (N = 42)	83%	10%	7%

FOEE= First Order Empirical Expectation.
FONE= First Order Normative Expectations.

Table 3. Normative and empirical expectations

When FOEE and FONE are different it is possible to explain subjects' behaviour through…			
	FOEE	FONE	OTHER
BT T (N = 14) M (N = 8)	72% 100%	14% 0%	14% 0%
AT T (N = 11) M (N = 9)	64% 78%	27% 22%	9% 0%
OT T (N = 14) M (N = 21)	57% 71%	14% 19%	29% 10%

FOEE= First Order Empirical Expectation.
FONE= First Order Normative Expectations.

Table 4. Groups' Voted Rule by University x Treatment.

		Rule					
		15 – 15 – 15 – 15		18 – 18 – 18 – 6		20 – 20 – 20 – 0	
Trento	AT	88.9%	8/9	11.1%	1/9	0.0%	0/9
	OT	87.5%	7/8	12.5%	1/8	0.0%	0/8
Milano							
	AT	100.0%	9/9	0.0%	0/9	0.0%	0/9
	OT	92.9%	13/14	0.0%	0/14	7.1%	1/14

Table 5. First voted rule by Treatment.

	AT	OT
15-15-15-15	75% 54/72	70% 62/88
18-18-18-6 or 20-20-20-0	25% 18/72	30% 26/88

Table 6. Compliance by University x Treatment.

Trento	AT	44.4%	12/27 10 rule 15 - 2 rule 18
	OT	29.2%	7/24
	OT (Insiders)	37.5%	6/16 5 rule 15 - 1 rule 18
	OT (Outsiders)	12.5%	1/8 1 rule 15
Milano	AT	55.5%	15/27 15 rule 15
	OT	45.2%	19/42
	OT (Insiders)	39.3%	11/28 9 rule 15 - 2 rule 20
	OT (Outsiders)	57.1%	8/14 7 rule 15 - 1 rule 20

Table 7. Distribution of FOEE by University x Treatment

		15 - 18	20
Trento	BT (N = 27)	15.0%	85.0%
	AT (N = 27)	20.0%	80.0%
Milano	BT (N = 15)	52.0%	48.0%
	AT (N = 27)	69.0%	31.0%

Table 8. Expectation of Compliance by University x Treatment.

Trento	AT	40.7%	11/27
	OT	20.8%	5/24
Milano	AT	51.8%	14/27
	OT	30.9%	13/42

Appendix 1– The beliefs elicitation procedure

Data on subjects' expectations were collected by means of a questionnaire. We used two different questionnaires, one for the Baseline and the Agreement treatments and one for the Outsider treatment.

BASELINE TREATMENT AND AGREEMENT TREATMENT

Let us identify the three active members of the group (players A) as Ax, Ay and Az. The questions were exactly the same for the three players. By way of example, we take the point of view of player Ax.

1. <u>First Order Empirical Expectations (FOEE)</u>

"You are participant Ax. According to you, what is the probability (expressed in percentage terms) that Ay has made the following choices:

CHOICE	PROBABILITY
S/he has asked for 15 tokens	[ ]
S/he has asked for 18 tokens	[ ]
S/he has asked for 20 tokens	[ ]

Remember that the three percentages must add up to 100%"

(We asked the subject if this probability would hold for player Az as well. If not, s/he could enter different values for Az. Thus, each subject answered two questions on FOEE.)

<u>2. Second Order Empirical Expectations (SOEE)</u>

"You are participant Ax. We now ask you to assign a probability (expressed in percentage terms) to each of these hypotheses regarding the probabilities assigned to your choice by participant Ay

HYPOTHESIS	PROB.
According to Ay, my most probable choice has been to ask for 15 tokens	[ ]
According to Ay, my most probable choice has been to ask for 18 tokens	[ ]
According to Ay, my most probable choice has been to ask for 20 tokens	[ ]

According to Ay, all my three choices are almost equiprobable [ ]

According to Ay, only two of my three choices are almost equiprobable [ ]

Remember that the five percentages must add up to 100%"

(We asked the subject if this probability would hold for player Az as well. If not, s/he could enter different values for Az. In this ways each subject were asked two questions on FOEE.)

3 First Order Normative Expectations (FONE)

"Think of a generic participant A. What is the right number of tokens that s/he should ask for?

I think the right number of tokens is 15 [ ]

I think the right number of tokens is 18 [ ]

I think the right number of tokens is 20 [ ] "

3 Second Order Normative Expectations (SONE)

"Think of a generic participant A. What do you think is her/his opinion about the right number of tokens that a generic participant A should ask for?

I think s/he believes that the right number of tokens is 15. [ ]

I think s/he believes that the right number of tokens is 18 [ ]

I think s/he believes that the right number of tokens is 20 [ ] "

OUTSIDER TREATMENT

In this treatment we have to distinguish between the members of the group who have voted for the rule and are still in their original group and the Outsider (the subject who come from a different group). We use "*Ax*" and "*Ay*" to denote the members who have not changed group and "*AO*" to denote the outsider.

1. First Order Empirical Expectations (FOEE)

Questions for the Ax and Ay members

"You are participant Ax (Ay). According to you, what is the probability (expressed in percentage terms) that Ay (Ax) has made the following choices:

(same options as in the other two treatments)

"You are participant Ax (Ay). According to you, what is the probability (expressed in percentage terms) that AO (the participant coming from another group) has made the following choices:

(same options as in the other two treatments)

Question for the AO members

"You are participant AO. According to you, what is the probability (expressed in percentage terms) that Ay (Ax) has made the following choices:

(same options as in the other two treatments)

2. Second Order Empirical Expectations (SOEE)

Questions for the Ax and Ay members

"You are participant Ax (Ay). We now ask you to assign a probability (expressed in percentage terms) to each of these hypotheses regarding the probabilities assigned to your choice by participant Ay(Ax).

(same options as in the other two treatments)

"You are participant Ax (Ay). We now ask you to assign a probability (expressed in percentage terms) to each of these hypotheses regarding the probabilities assigned to your choice by participant AO (the participant coming from another group):

(same options as in the other two treatments)

Question for the AO members

"You are participant AO. We now ask you to assign a probability (expressed in percentage terms) to each of these hypotheses regarding the probabilities assigned to your choice by participant Ax (Ay):

(same options as in the other two treatments)

3 First Order Normative Expectations (FONE)

Questions for the Ax, Ay and AO members

"Think of a generic participant A who is still in her/his original group. What is the right number of tokens that s/he should ask for? **(FONE1)**

(same options as in the other two treatments)

"Think of a generic participant A who is in a group which is not her/his original one. What is the right amount of tokens that she/he should ask for? **(FONE2)**

(same options as in the other two treatments)

4 Second Order Normative Expectations (SONE)

Questions for the AO members

"Think of a generic participant A who is still in her/his original group . What do you think is her/his opinion with regard to the right number of tokens that a participant A who is still in her/his original group should ask for ?" **(SONE1)**

(same options as in the other two treatments)

"Think of a generic participant A who is still in her/his original group . What do you think is her/ his opinion with regard to the right number of tokens that a participant A who is not in her/his original group should ask for ?"

(SONE2)
(same options as in the other two treatments)

Questions for the Ax and Ay members

Think of a participant A who is still in her/his original group . What do you think is her/his opinion with regard to the right number of tokens that a participant A who is still in her/his original group should ask for ?
(SONE3)

(same options as in the other two treatments)

Think of a participant A who is still in her/his original group . What do you think is her/his opinion with regard to the right number of tokens that a participant A who is not in her/his original group should ask for ?
(SONE4)
(same options as in the other two treatments)

"Think of a participant A who is not in her/his original group . What do you think is her/his opinion of the other participant A with regard to the right number of tokens that a participant A who is still in her/his original group should ask for ?"
(SONE5)

(same options as in the other two treatments)

Think of a participant A who is not in her/his original group . What do you think is her/his opinion of the other participant A with regard to the right number of tokens that a participant A who is not in her/his original group should ask for ?
(SONE6)
(same options as in the other two treatments)

Subjects were paid only for the accuracy of their guesses in FOEE and SOEE questions according the Quadratic Scoring Rule (Davis and Holt, 1993).

When we examine the relation between subjects' choices and beliefs, we consider only first order expectations (both empirical and normative). This is due to a preliminary analysis on beliefs. We first analyse First Order Empirical Expectations (FOEE) and Second Order Empirical Expectations (SOEE). In particular, we want to check whether what subjects think the others have done is in line with what they think the others expect s/he had done. We find that

there is no difference between FOEE and SOEE in all the treatments ($p < 0.06$, Fisher-exact test in the BT; $p > 0.45$, Wilcoxon test in the AT; $p > 0.15$, Wilcoxon test in the OT).[12]

We then check whether this is also the case when considering normative expectations. In the BT, it turns out that First Order Normative Expectations (FONE) and Second Order Normative Expectations SONE are not significantly different ($p = 0.000$, Fisher-exact test). In the AT, FONE are slightly lower than SONE ($p = 0.09$, Wilcoxon test), but highly correlated ($p = 0.0002$, Spearman correlation test). In the OT the analysis is rather more complicated because we have two different kinds of active players – the outsiders and the insiders. Consequently, normative beliefs concern both a generic insider and a generic outsider rather than a generic player A – as in the BT and in the AT. This increase the number of normative expectations (FONE1, FONE2, SONE1, SONE2, SONE3, SONE4, SONE5 and SONE6) and the number of possible comparisons. With respect to the outsiders, we compare FONE1 with SONE1 and FONE2 with SONE2. We find that FONE1 and SONE1 are not significantly different ($p = 0.34$, Wilcoxon test), while FONE2 are slightly lower than SONE2 ($p = 0.05$, Wilcoxon test). However, when we compared SONE2 with choices, it turns out that they are not significantly correlated ($p = 0.41$, Spearman correlation test). Concerning the insiders, we compared FONE1 with SONE2 and SONE5, as well as FONE2 with SONE4 and with SONE6. In all cases we find that they are not significantly different ($p > 0.31$, Wilcoxon test). Finally, we check whether players think that a normative choice does not depend on the role (outsider vs insider). We compare FONE1 with FONE2 and found out that they are not significantly different for both the outsiders ($p = 0.34$, Wilcoxon test) and the insiders ($p = 0.19$, Wilcoxon test).

To sum up, we find that second order expectations are generally in line with first order expectations. This makes it to study the relation between choices and beliefs by taking only first order expectations into account.

[12] We would stress that when the tests are run, the independence of observations is taken into account. In particular, in the BT each player's observation is independent from all the other players' observations. In the AT, independent observations are the group's average observations. In the OT, the insiders' independent observations are again the group's average observations, while the outsiders' independent observations are the average observations of the interchanged outsiders.

Appendix 2 – The Econometric Analysis[13]

$$FOEE\_20_i = AT_i\alpha + Age_i\beta_1 + TENT_i\beta_2 + FIRST^* AT_i\delta_1 + \varepsilon_{i1} \quad \text{(R1)}$$

(R1) is a probit regression that we implement to explore what variables influence the subjects' probability of expecting that the others have chosen 20. The dependent variable is the dichotomous variable FOEE_20, which is equal to 1 if a subject expects the others to have chosen 20. The control variables are both related to the nature of the experiment (AT, FIRST*AT, TENT) and demographic (AGE). We exclude the variable GENDER since it turns out that in the first two treatments GENDER and AGE are significantly correlated (Pearson coefficient; $p < 0.01$) – the women are significantly older than the men (ttest; $p = 0.002$). AT is a dummy equal to 1 if the AT is played. TENT is the number of voting rounds for the group before it reaches a unanimous decision on the rule to be used – variable equal to 0 when the BT is played. FIRST*AT is an interaction term equal to 0 either when the BT is played or when the player in the AT have participated in other experiments in the past.

Probit Model – R1

Variables	***FOEE_20***	***Marginal Effects***
AT	-2.1***	-0.58
	(0.478)	
FIRST*AT	-1.29***	-0.47
	(0.453)	
AGE	-0.10	-0.03
	(0.073)	
TENT	0.39**	0.13
	(0.169)	
Constant	3.77***	
	(1.643)	

N	96
Log Likelihood	-39891664
LR chi2(4)	42.43
Prob > chi2	0.000

*****significance 1%**

**** significance 5%**

[13] Multicollinearity – a usual problem of probit regressions – is detected using VIF tests.

(R1) show that subjects are more likely to expect selfish behaviour by the co-players in the BT. Moreover, it emerges that, in the AT, the higher the number of voting rounds before the group reaches a unanimous decision on the rule to be used, the higher the probability of the subjects expecting selfish behaviour by the co-players. Finally, in the AT, a player who has never participated in other experiments in the past has a higher probability of asking for a sum different from 20.

$$FOEE\_20_i = AT_i\alpha + FIRST^* AT_i\beta_1 + TENT_i\beta_2 + AGE_i\beta_3 + \varepsilon_{i1}$$
$$CHOICE\_20_i = FOEE\_20_i\delta_1 + FIRST^* AT_i\beta_4 + AGE_i\beta_5 + \varepsilon_{i2} \quad \text{(R2)}$$

(R2) is a bivariate recursive probit regression[14] where *CHOICE_20* is equal to 1 if subject *i* choose 20 tokens. It makes it possible to check: 1) the relation among agreement, beliefs and choices; 2) whether there is any latent variable that might influence beliefs and choices at the same time.

Bivariate Recursive Probit Model – R2

Variables	***FOEE_20***	***CHOICE_20***
AT	-2.87***	
	(0.57)	
FIRST*AT	-1.4***	-0.04
	(0.422)	(0.433)
AGE	-0.15*	0.11
	(0.085)	(0.095)
TENT	0.40**	
	(0.168)	
FOEE_20		2.42***
		(0.712)
Constant	8.16***	-4.38*
	(2.3)	(2.365)

N	96
Log Likelihood	-73.623096
Rho	0.287
Prob > chi2	0.47

*****significance 1% ** significance 5% * significance 10%**

[14] A variation of the analysis run by Di Novi (2007).

(R2) shows that the agreement influences empirical expectations and that empirical expectations influence the subjects' decisions. Moreover, because rho is not significantly different from 0, we can state that there is no latent variable influencing beliefs and choice at the same time.

$$EQFOEE_i = OT_i\omega + TENT_i\phi_1 + FIRST_i\phi_2 + \nu_{i1} \quad \text{(R3)}$$

(R3) is a probit regression that we implement to explore what variables influence the subjects' probability of expecting that the others have chosen the voted rule. The dependent variable is the dichotomous variable EQFOEE, which is equal to 1 if a subject expect the others to have chosen the voted rule. The control variables are all related to the nature of the experiment (FIRST and TENT). We exclude all demographic variables because there is no significant difference due to gender (chi2; $p = 0.97$) and the variables AGE and FIRST are significantly correlated (Pearson coefficient; $p < 0.05$).

Probit Model - R3

Variables	***EQFOEE***	***Marginal Effects***
OT	-0.48**	-0.18
	(0.242)	
FIRST	0.32	0.118
	(0.247)	
TENT	-0.09	-0.03
	(0.069)	
Constant	0.01	
	(0.253)	
N	120	
Log Likelihood	-74.073703	
LR chi2(3)	8.44	
Prob > chi2	0.0539	

**** significance 5%**

(R3) shows that subjects are more likely to expect compliance by the co-players in the AT.

$$EQFOEE_i = OT_i\omega + TENT_{i1}\phi_1 + FIRST_{i1}\phi_2 + v_{i1}$$
$$EQCHOICE_i = EQFOEE_i\delta_2 + AGE_i\phi_3 + v_{i2} \quad \text{(R4)}$$

As in the comparison between the BT and the AT, we compare the AT and the OT by running a bivariate recursive probit (R4) where EQCHOICE is equal to 1 if the choice corresponds to the voted rule.

Bivariate Recursive Probit Model – R4

Variables	*EQFOEE*	*EQCHOICE*
OT	-0.47**	
	(0.243)	
FIRST	0.40	
	(0.27)	
AGE		0.05
		(0.057)
TENT	-0.07	
	(0.092)	
EQFOEE		2.39***
		(0.945)
Constant	-0.09	-2.065
	(0.342)	(1.284)

N	120
Log Likelihood	-133.37077
Rho	-0.51
Prob > chi2	0.579

*****significance 1%**

**** significance 5%**

(R4) shows that introduction of the mixed protocol influences empirical expectations and that empirical expectations influenced the subjects' decisions. Moreover, because rho is not significantly different from 0, we can state that there is no latent variable influencing beliefs and choices at the same time.