Other-regarding decisions: Fairness and normative expectations in a doctor-patient experiment

Timo Tammi
Department of Economics and Business Administration,
University of Joensuu, Finland. E-mail: timo.tammi@joensuu.fi.
Ismo Linnosmaa
Department of Health Policy and Management,
University of Kuopio, Finland
Sami Pakarinen
Department of Economics and Business Administration,
University of Joensuu, Finland
Other-regarding decisions: Fairness and normative expectations in a doctor-patient experiment

Timo Tammi
Department of Economics and Business Administration,
University of Joensuu, Finland. E-mail: timo.tammi@joensuu.fi.
Ismo Linnosmaa
Department of Health Policy and Management,
University of Kuopio, Finland
Sami Pakarinen
Department of Economics and Business Administration,
University of Joensuu, Finland

Abstract

Behavioral and experimental economists have recently put much effort on analyzing non-selfish behavior and the interplay between selfish, other-regarding, inequality aversive and altruistic motivations of action. Most of the experiments where these questions are investigated are based on the idea of experiments as abstract (or generic) contexts of interpersonal behavior. In this paper we take another strategy by focusing on a situation with a specific context and cues that direct the players’ attention to normative and empirical expectations relevant to the situation. The experimental analysis is based on the so called doctor-patient game, where subjects, acting in the role of a doctor, choose between an expensive treatment and an inexpensive one. The expensive treatment gives a low payoff to the doctor but a high payoff to the patient; the inexpensive treatment gives a high payoff to the doctor but a low payoff to the patient. Simultaneously with the doctor’s choice the patient chooses between complying or not complying with the doctor’s decision. We found that the majority of the doctor’s choices deviate from the prediction of the traditional theory and that the choices are influenced by the doctor’s second order empirical expectations and by the doctor’s outcome fairness motivation concerning the payoff-difference in the game.

Keywords: other-regarding preferences, social norms, behavioural economics, experimental economics
JEL classification: C7, C9

1 Participants in seminars at the University of Joensuu in February 2009 and in December 2009 and participants in the Economic Science Association (ESA) European Conference in Innsbruck in September 2009 are acknowledged for constructive discussions. The work has been supported by Yrjö Jahnsson Foundation.
1. Introduction

A positive other-regarding decision occurs when a person sacrifices her material gain in order to generate a benefit to some other person. Such other-regarding decisions abound in all human life, an obvious exception being anonymous interaction in competitive markets. Accordingly, we see that people leave tips in restaurants they never visit again, pay for music and games they could download free of charge, take risks when not firing employees during an economic decline, contribute to common projects, are satisfied with lower-than-competitive level wages when working in the non-profit sector, and so on. Furthermore, spontaneous behavioural regularities and norms have emerged as well as many legal rules have been designed to advance other-regarding behavior. And, in branches such as health, welfare and legal services, for example, professional experts are expected to promote the ends of their clients.

In spite of their predominant application of the assumptions of rationality and self-interest (Zizzo 2003; Sobel 2005) economists have long known that people care of others and of what others think of them. However, it was not until recently that economists became interested in other-regarding behaviour in a larger scale. To date, hundreds of experiments by economists as well as analyses of field-evidence have shown that in various situations of personal interaction a remarkable portion of subjects are not motivated merely by their own material gains. That is to say, things such as others’ material gains, intentions of others, aversion of inequality etc. often play a major part in subject's decisions. A recent topic is social norms and norm-related expectations as well as the influence of norms and expectations on the behaviour of subjects in various economic settings.

In this paper we employ the (wide interpretation of the) framework of psychological game theory to analyze other-regarding decisions in a situation we call a doctor-

---

We here refer to the ‘pay-what-you-want’ pricing experiments in the internet delivery of the rock band Radiohead’s album in Autumn 2007 and the video game company 2D Boy’s World of Goo game in Autumn 2009.
patient experiment. Psychological game theory suggests that in a situation involving personal interaction, like that between a doctor and a patient in our experiment, utilities of a player depend not only on her material outcomes, but also on the beliefs of others – especially on the second order beliefs of the opponent player(s). In the doctor-patient experiment, therefore, we expect that the doctor’s decision is influenced significantly by what the doctor believes is the normal practice (and its prevalence) among real practicing doctors and by what the patient expects the doctor to decide. Regarding these expectations, we elaborate the distinction between two genres of second order expectations – namely, those based on one’s view to population distribution and the others formed on the basis of one’s model of the internal state of the opposite side of the interaction. In addition, we look out for a fairness-effect, viz., whether the increase in the payoff-difference between the doctor’s and the patient’s material outcomes fortifies the behavior compliant with the normative expectations.

The doctor-patient experiment can also be seen as a stylized account of interaction between a professional specialist and a client. Both actors have material incentives which may or may not be in a conflict with the specialist’s motivation to comply with the normal practices (and the ethical standards behind them) of his/her profession (see Sacconi 2007) and with the client’s expectation of the specialist’s behavior. To account for such interplay we construct a model of normative expectations and fairness. The model utilizes the insights presented particularly in Sugden’s (1998; 2000) theory of normative expectations but to a smaller degree also in Fehr and Schmidt’s (1999) inequality-aversion theory and in the quilt-aversion model by Charness and Dufwenberg (2006) and Battigalli and Dufwenberg (2007).

In order to examine the interplay of material incentives and normative expectations, we develop a doctor-patient dictator game (henceforth DPD game) and a doctor-patient treatment game (henceforth DPT game). In our experimental design the DPD game is effective in identifying outcome fairness in a general medical context. In the game, a subject, when acting in the role of a doctor, is asked to divide a fixed budget between her own salary and the treatment cost to a patient of which the doctor only
knows that a larger sum allocated to the patient gives a higher health benefit to the patient. We suggest that many of those who allocate the budget on the fifty-fifty basis exhibit both outcome-fairness motivation and compliance with the normal practices among real practicing doctors. This knowledge helps us in analyzing the significance of the increase of the payoff-difference in the DPT game.

In the DPT game, the doctor selects between an expensive and inexpensive treatment. Simultaneously with the doctor’s decision, the patient chooses whether or not to comply with the doctor’s treatment choice. A utility function with only selfish preferences implies that the doctor in the DPT game chooses the inexpensive treatment while a utility function with strong enough other-regarding preferences would imply the choice of the expensive treatment. We suggest, however, that the fact that the majority of the subjects choose the expensive treatment need not result from other-regarding preferences but from normative expectations related to the ethical standards of medical decision-making.

The paper is organized as follows. We first discuss in section 2 on the role of norms and expectations in economic behaviour and then in section 3 we present our model of normative expectations and fairness. Section 4 presents the doctor-patient experiment and the predictions of our model in the context of that experiment. Section 5 deals with the results of the experiment and in section 6 we gather up the findings of the study in the context of relevant earlier findings. Section 7 concludes the paper.

2. Norms, normative expectations and economic behavior

2.1. Emotional sanctioning model

A norm is here defined as a regularity of behavior shared and known by a group of people. Economists typically see that a norm is external to a person: a person conforms with a norm whenever it is a good bet to do so. Sociologists, on the other hand, tend to see that people comply with norms because they have internalized them
(see Anderson 2000). We here take the economists’ point of view but enrich it with some psychological constituents. In effect, we are interested in providing theoretical and empirical evidence that some behavioral regularities explained usually in terms of other-regarding preferences can be understood, at least partly, as effects of the motivational power of norm-related expectations.

Out of several theoretical approaches which combine rational choice with social norms we here focus on an emotional sanctioning model. According to this model, the sanctions related to norms are ‘costless’ and ‘automatic’ and ‘unintended’ as they are emotional/psychological tendencies to seek approval and to avoid disapproval. The model is in tune with some evidence from cognitive psychology of thinking and neuroscientific analysis of norm compliance.

According to the emotional sanctioning model, norms as regularities (R) bring about normative expectations in the following way (Sugden 2000, 82):

(1) if a person conforms to R, it is her interest that others also conform to R;
(2) in most cases, if a person expects others to conform to R, it is her interest to conform to R, too; and
(3) most people in society expect most other people to conform to R.

In referring to the motivating power of norm-compliance Sugden further introduces the notion of resentment-aversion. Accordingly, people are resentment-aversive in the sense that those who conform to normative expectations will feel resentment against those who do not conform. If this is ‘common knowledge’ then a person may be

---

3 See Anderson (2000) and Guala and Mittone (2009) for an analysis of the different approaches.
4 The so called dual process theory of thinking provides an account of how a person categorizes a situation and how a specific norm becomes salient (see Payne and Bettman 2004). Some recent studies on behavioral and neural correlates of norm compliance show that brains are well ‘wired’ to cope with emotional aspects of norm compliance and the related possibilities of punishing and rewarding (Spitzer et al. 2007). An essential role is in an individual’s ability to model the other player’s internal state. This provides a basis for an individual to form expectations of the other player and of her expectations in addition to her view of the distribution in the general population (see Rustichini 2005).
motivated to avoid other’s resentment against her and believe that the others may also be motivated to avoid her resentment towards them.

On the other hand, Hargreaves Heap and Varoufakis (2004) suggest that human beings are not only resentment-averse but also have (what they call) subversive tendency – viz. propensity towards overthrowing others’ expectations. Furthermore, this is only one form of non-compliance to a norm. Another case is the co-existence of rival norms and corresponding expectations: a person may deviate from one norm (R1) but comply with the other norm (R2) if she either (i) is motivated to frustrate the expectations of those whose interest is everybody’s compliance with R1 or (ii) she is motivated to comply with the expectations of those whose interest is everybody’s compliance with R2. Although the question of non-compliance cannot be further discussed here, it is important to note that an analysis of the reasons of a person’s non-compliance is needed in explaining the observed heterogeneity of behaviours in experiments.

2.2. Some recent views to norm-compliance in experiments

A dictator game is a behavioural ‘game’ which has been studied in hundreds of experiments by economists since 1986 (Kahneman et al. 1986). In a dictator game, a player called ‘the proposer’ allocates some money amount to herself and to a player called ‘the responder’. The average allocation to the responder is 25% of the whole amount but the allocation varies considerably from 13% to 50% (Camerer 2003, 56-59) or even to 72% (Brañas-Garza 2006). The received view has been that the non-negative, often considerable allocations to the responder are results of other-regarding preferences – that is, the proposer’s utility is an increasing function of her own and the respondent’s material payoffs.

List (2007) and Fershtman et al. (2008) reinterpret dictator game results along the lines of some unspecified model of norm-induced behaviour. List found that when the dictator game setting is varied to weaken the saliency of the norm of ‘equal split’ then
fewer subjects behave generously. List further argues that rather than referring to other-regarding preferences the data results from the expectations the players have – possibly influenced by social norms (List, 490). In a similar vein, Fershtman et al. (2008) found that in dictator games and trust games the properties of the situation determined the social norms which defined the socially accepted choices. In particular, they found that when the baseline dictator game environment (where presumably the equal split norm is the salient one) was changed into a more competitive environment (where, as the authors suggest, emerges a norm that legitimizes unequal partition), then the frequency of unequal splits also increased at the expense of equal divisions.

Bicchieri and Chavez (2009) and Bicchieri and Xiao (2009) also report on the influence of social norms and norm related expectations on the behavior of subjects in dictator game and ultimatum game experiments. These papers make use of the distinction between empirical expectations (what we expect others to do) and normative expectations (what we believe others think we ought to do). It seems that expectations have a crucial role in subject behaviour: both normative and empirical expectations predict one’s behavior but when empirical and normative expectations are in conflict, empirical expectations have a decisive role.

Faillo et al. (2008) and Sacconi and Faillo (2008) study the norm complying behavior in situations comparable to dictator game. They (Faillo et al., 2008) also found that subjects behaved according to the salient empirical and normative expectations, and, when empirical and normative expectations were in conflict, according to the empirical ones. Their experiment differed from the dictator game but the results allow a comparison of the proportions of selfish and equal allocations across different environments. Especially, when an equal division norm was made salient via voting among different rules of division, the proportion of selfish choices diminished. Sacconi and Faillo (2008) refer to John Rawls’s idea of the sense of justice in explaining subjects’ behavior. To put it shortly, the interpretation suggests that we learn a sense of justice during our lives and are therefore ‘souped-up’ to replicate
conformity with a publicly known just norm, on the condition we believe that others are similarly endowed and disposed.

2.3. On second order expectations

Although we follow List (2007), Freshtman et al. (2008), Bicchieri and Xiao (2009), and Faillo et al. (2008) in analysing norm-related expectations in general, we introduce a distinction between (i) a player’s expectation as a hypothesis about the distribution of norm-compliance in (a relevant) population and (ii) a player i’s guess of the opposite player’s internal state and her corresponding expectation of i’s choice. We categorize the first kind of expectation as an indirect second order empirical expectation (ISE) and the second kind of expectation as a direct second order empirical expectation (DSE). We take DSE as an unproblematic case (at least presently) but provide a short elaboration of ISE.

In elaborating ISE we make use of Sugden’s (2000) resentment hypothesis. Accordingly, assume two players d (doctors) and t (patients): players d and t form a population P and they interact in situations of type S, where d has two alternative actions x and y. Every time d makes her decision, it will become common knowledge among P. Let us now suppose that d has normally chosen x. It has therefore become common knowledge that d in S normally chooses x. Because of this common knowledge R, t has good reasons to form an empirical expectation that d will choose x and not y. This is also known by d: she knows that t expects that she chooses x. Then, since there is the commonly known regularity R, d’s choice of y would evoke resentment in t towards d. Since also d realizes this, d is aversive towards choosing y. We can now rephrase ISE and DSE as follows:

[ISE]: d’s and t’s hypotheses about the distribution of norm-compliant behavior among ds in situations of type S;
[DSE]: d’s guess of t’s internal state (e.g., t’s willingness to comply with d’s choice) and t’s corresponding expectation of d’s choice in situations of type S.
We further suggest, that both ISE and DSE have power to influence the decisions of doctors and that they are often correlated. However, DSE is sensitive to the clarity of its internal state while ISE is not.

In the next section, where we present the theoretical model that takes into account the second order expectations, we omit the above distinction. Then again in section 4 we come back to the distinction between ISE and DSE by defining them as empirically measurable variables. The idea is, that in the context of our doctor-patient experiment, ISE is measured by asking subjects to assess the prevalence of practicing doctors’ compliance with the normal practices of medical decision-making. If, for example, a subject believes that about 90% of doctors comply with the normal practices, then her indirect second order expectation is that patients see that the doctors very probably make a decision compliant with the normal practices. DSE, on the other hand, is measured by asking subjects to assess, in the role of a doctor, directly, what they think is the probability the patient attaches to the possibility that the doctor chooses a particular action in the game. The experimental treatment effect related to this comes from the difference in how the patient’s self-care efforts are associated with the patient’s decision to comply with the doctor’s treatment decision. The patient’s inner state is transparent in the baseline case where higher self-care effort means higher health benefit to the patient. On the contrary, the inner state is opaque in the modified case where higher self-care effort means lower health benefit to the patient. 

3. The model

The model we present in this section is based on the idea that a person feels utility both from her own material outcomes and, possibly, also from the fairness of the distribution of the outcomes and from her own compliance with the empirical

5 On transparency and opaqueness, see Tversky and Kahneman 1986, Hogarth and Reder (1986), and Irwin et al. (1998). In general, the opaqueness of a decision-making situation (or, a part of it) increases the cost of thinking and emboldens the decision-maker to consult additional devices of making up her/his mind. Hunches, authority, norms, rules-of thumb and next-best preferences are amongst the popular devices.
expectations salient in the situation. We also lean on the idea of the variable frames theory of games (Bacharach and Bernasconi 1997) in that different players may play the game differently depending on the frame they use in analyzing and deciding in the game.

Let us first assume an abstract pure strategy game, where the row-player (R) chooses between Top and Bottom strategies (T and B, respectively) whereas the column-player chooses between Left and Right strategies (L and R, respectively). Let $a, b, c, d$ and $e$ denote the payoffs of the game shown in figure 1. We assume that $a > b > c > d > e$.

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>$b, b$</td>
<td>$b, d$</td>
</tr>
<tr>
<td>B</td>
<td>$a, c$</td>
<td>$a, e$</td>
</tr>
</tbody>
</table>

*Figure 1. The abstract game*

We can see that the pure-strategy Nash equilibrium in the game is the strategy pair (B, L) with corresponding pay-offs $a$ and $c$. Hence, the row-player always chooses ‘bottom’ and the column-player chooses ‘left’.

It is now interesting to consider how real-world people might play the game. Do they find the same solution as the game theorists? Presumably not, as real world people are often experienced in solving problems in contextual games whereas economists typically are experts in abstract games. Let us bring a minimal amount of context by presenting the game in a specific terms according to the rule $a > b > c > d > e$. The game with numbers, or, if you like, in Euros, is shown in figure 2.

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>6, 6</td>
<td>6, 3</td>
</tr>
<tr>
<td>B</td>
<td>7, 5</td>
<td>7, 2</td>
</tr>
</tbody>
</table>

*Figure 2. The abstract game with monetary pay-offs*
Obviously, most column-players may now reason that it is always better to choose ‘L’ and that the final outcome depends on the row-player’s decision. Most of the row-players, on the other hand, might very likely think that one would give one Euro’s benefit to the column-player at one Euro’s cost to herself. The actual choice of the row-player depends on her motivation: if she wants to maximize her own material pay-off, she chooses ‘B’ but if, for example, she is a ‘pure altruist’, she would rather choose ‘T’.

Imagine then that the game involves a behavioural regularity, introduced to the players, that in this kind of a game, the row-player is commonly observed to choose ‘T’. This brings normative and empirical expectations into the minds of the players. R knows that C has expectations regarding R’s choice and that these expectations are shaped by the regularity just introduced. It is now possible, that C feels disappointment if R deviates from C’s expectations and that R knows this.

In a real-world situation the context as well as the participant’s motivation and available cognitive resources determine which of the feasible regularities become(s) salient. Regarding experiments, the labelling of alternative strategies as well as the role of players is a means to manipulate the context and its salient features. In order to analyse a situation of personal interaction where normative expectations are present, we next construct a model that incorporates self-interest, inequality-aversion and expectations. Such a utility function takes the following shape:

\[ V_i = \pi_i(s) + \beta_i^b \times \lambda_i M(\pi_j(s), b_j^2), \]

where \( V_i \) is \( i \)'s comprehensive utility resulting directly from \( i \)'s own monetary gains and from \( i \)'s compliance with or deviation from \( j \)'s expectations regarding \( i \)'s behaviour. The parameters \( \lambda \) and \( \beta \) are weights of the dis-utility resulting from the deviation from the expectations. More precisely, \( \lambda_i \geq 0 \) stands for the importance \( i \)

---

6 See Bacharach and Bernasconi (1997) and Gold and Sugden (2006), and, in general, Schelling (1960). In particular, the variable frame theory by Bacharach would suggest, that the frame the player uses, influence on what she decides in the game.
places on others’ disapproval in general (NB: when $i$’s choice is the one implied by the norm, then $\lambda_i = 0$, which means that in such a case $i$’s utility results only from her own monetary gains). On the other hand, $\beta_i$ links $\lambda M(\cdot, \cdot)$ with the structural payoff inequality in the game. Formally,

$$ (2) \quad \beta_i = 1/\left(\frac{|\overline{\pi}_i|}{|\overline{\pi}_j|}\right) $$

where $\overline{\pi}_j \neq 0$ and $\overline{\pi}_i \neq 0$ are means of $j$’s and $i$’s pay-offs, respectively. The exponent $b$ in equation (1) is a binary exponent: $b = 0$ when the player has no outcome-fairness motivation and $b = 1$ when she has. The $\beta_i$ approximates the structural inequality of the game from $i$’s point of view. We can say that structural equality prevails if $|\overline{\pi}_j| = |\overline{\pi}_i|$ and $\beta_i = 1$. When $|\overline{\pi}_j| < |\overline{\pi}_i|$ and $\beta_i > 1$, the impact of favourable structural inequality (from $i$’s point of view) increases the magnitude by which $i$’s deviation from normative expectations decreases $i$’s utility. On the other hand, when $|\overline{\pi}_j| > |\overline{\pi}_i|$ and, consequently, $\beta_i < 1$, there is unfavourable structural inequality from $i$’s point of view. This weakens the negative impact of $i$’s deviation from the normative expectations.

Finally, let us concretise the function $M(\pi_j(s), b^2_j)$. The function says that the utility related to normative expectations depends on the payoffs of $j$ and on $i$’s second order expectations $b^2_j$. A simple formulation goes as follows:

$$ (3) \quad M(\pi_j(s), b^2_j) = \pi^*_j - E(\pi_j), $$

where $\pi^*_j$ is $j$’s actual pay-off chosen by $i$, and $E(\pi_j)$ is $j$’s expected pay-off seen in the eyes of $i$ and based on the probability $p$ $i$ assigns to $j$’s expectation that $i$ chooses

---

7 The idea here is similar to that in Sugden (2000) and in Battigalli and Dufwenberg (2007). Thus, $M(\cdot, \cdot)$ measures the disutility one feels if she chooses so as to give to the other party less than she expects the other party to expect.
a particular strategy. Note also that $M(.,.)$ is restricted to negative values only: if 
$\pi_j - E(\pi_j) \geq 0$, then $M(.,.)$ is zero$^8$. The following predictions can be deduced from 
the model with reference to the game presented above:

(1) In an abstract situation, where the players’ roles are labelled abstractly (‘R’ 
and ‘C’) and where players do not know each others, most of the row-players 
choose B and column-players choose L;

(2) When a social context, that implies the desirability of the strategy-pair (T, L) is 
introduced into the game (by re-labelling the roles and/or by describing the 
situation), most row-players choose T and column-players choose L;

(3) The introduction of the social context generates expectations among row-
players, that most column-players assign a high probability to the alternative that 
the row-player chooses T;

(4) If the payoff-difference, which is advantageous to R, increases, more row-
players shift from B to T.

4. The doctor-patient experiment

4.1. The design of the experiment

The doctor-patient experiment was conducted at the University of Joensuu on 
February 10, 2009 and at the University of Kuopio on February 12, 2009. The 
subjects in the experiments were undergraduate students from both universities. They 
were recruited via an e-mail announcement to students who have been registered for 
the general subject pool for experiments at the two universities. The total number of 
recruited subjects was 56 (30 in Joensuu and 26 in Kuopio). Of the subjects 12,4% 
were students of economics or business economics, 37,5% of social sciences and 
humanities, and 50% of health sciences, medicine and natural sciences.

$^8$The rationale of this is in the fact that the model accounts for only such deviances from the other 
party’s expectations that yield to her a smaller pay-off than she expected.
The experiment consisted of three stages and a final phase where the subjects filled a survey concerning demographic and other background information. Two stages were played for real money by using ‘points’ as an experimental toy-money. The participants were told in the beginning of the experiment and in the written guidelines that 250 points was worth of 1 Euro. The average earning in the experiment was 15,10€.

The experiment took place in two rooms simultaneously. The participants were divided between the rooms in a random way. In the first room the experiment progressed in the following way. In stage one each subject made her decision in a doctor-patient dictator game task – that is, she allocated 1000 points between herself (in the role of a physician) and an anonymous receiver (in the role of a patient).

The stage two consisted of two rounds. In the first round the subject made a decision to choose either an expensive treatment E or an inexpensive treatment C in a situation described in game G1 (table 2). In the second round the subject made a similar decision in a situation described in game G2 (table 2). In the third stage a subject made her decisions to play either ‘high-level self-care effort’ H or ‘low-level self-care effort’ L in both situations game G1 and game G2. As can be seen, there are two differences between G1 and G2: (1) the mean payoff-difference is greater in G2 than in G1, and (2) the patient’s internal state (reasoning of how her health benefit results from high or low self-care effort) is clearer (transparent) in G1 but less clear (opaque) in G2. Regarding the second difference, we mean that it is generally easier to grasp that high self-care effort is connected to the greater payoff than that low self-care effort is connected to the greater payoff.

Table 2. The doctor-patient games G1 and G2

<table>
<thead>
<tr>
<th></th>
<th>G1</th>
<th></th>
<th>G2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patient</td>
<td></td>
<td>Patient</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High (H)</td>
<td>Low (L)</td>
<td>High (H)</td>
<td>Low (L)</td>
</tr>
<tr>
<td>Doctor</td>
<td>Expensive (E)</td>
<td>600, 600</td>
<td>Expensive (E)</td>
<td>600, 200</td>
</tr>
<tr>
<td></td>
<td>Cheap (C)</td>
<td>700, 500</td>
<td>Cheap (C)</td>
<td>700, 100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>600, 300</td>
<td></td>
<td>600, 300</td>
</tr>
<tr>
<td></td>
<td></td>
<td>700, 200</td>
<td></td>
<td>700, 200</td>
</tr>
</tbody>
</table>
In the second room the order of stages was the reversal of the order in the first room – that is, beginning from stage three, followed by stages two and one, and ending up with the final phase. Only negligible order effects were found.

The final phase⁹ in the experiment the subjects were asked to recall his/her decisions as a doctor and to (A1) express the probability she/he thought that the patient assigned to her/his decision to choose the expensive treatment in G1 and in G2. In addition, the subjects were asked to express her/his view of the following norm that we think describes the ethical standard in medical decision-making: *the doctor should always and only pay attention to the patient’s benefit in prescribing a treatment to the patient*. The subjects were asked to (A2) estimate which percentage of real doctors complies with the norm. In the data-analysis we use A1 as a yardstick of direct second order expectations (DSE) and A2 as a yardstick of indirect second order empirical expectation (ISE) defined in section 2.3.

4.2 Model’s predictions in the doctor-patient game context

To derive the predictions of our theoretical model in the doctor-patient treatment (DPT) game we first produce a new description of the game by adjusting payoffs according to the model. Since the doctor’s payoff in the game does not depend on the patient’s choice, the function $M_p(.,.)$ in the patient’s comprehensive utility function $V_p$ gets the value of zero. For example, in the game G1, when the doctor’s choice is E and the patient’s choice is L, $M(.,.) = \pi_d(s) – E(\pi_d) = 600 – 600 = 0$. Recall that we also defined that $\lambda = 0$ if the player complies with the normative expectations: in that case, no disapproval is aroused. Since $\lambda_d = 0$ in the cases (E, H) and (E, L) we get the following utility functions that bring about the adjusted payoffs in game G1, where $\beta_d$,

---

⁹ This phase consisted also of a questionnaire of background information on demographic factors and the so called standard altruism scale questions (see Rushton et al. 1981). There were also some additional questions (i) of expectations of the probability of the patient’s high self-care effort, (ii) of the subject’s compliance with the norm if she/he were a real doctor, and (iii) a control question of the subject’s beliefs of which percentage of real doctors complies with the norm. The after-experiment phase also included a hypothetical strategic decision in a two-person game known as ‘relying to the rationality of the other’ (Rubinstein 1999).
the measure of advantageous structural payoff-difference for the doctor, gets the value of 1,6:

\[ V_d(C, H) = 700 + 1.6^b \lambda [500 - (p600 + (1 - p)500)] = 700 - 1.6^b \lambda 100p. \]

\[ V_d(C, L) = 700 + 1.6^b \lambda [200 - (p300 + (1 - p)200)] = 700 - 1.6^b \lambda 100p. \]

In game G2, where \( \beta_d \) increases to the value of 3,3, the corresponding utilities are:

\[ W_d(C, H) = 700 + 3.3^b \lambda [100 - (q200 + (1 - q)100)] = 700 - 3.3^b \lambda 100q. \]

\[ W_d(C, L) = 700 + 3.3^b \lambda [200 - (q300 + (1 - q)200)] = 700 - 3.3^b \lambda 100q. \]

In the above equations \( V_d \) and \( W_d \) refer to the comprehensive utilities of the doctors, and \( p \) and \( q \) denote to the probabilities (in G1 and G2, respectively) the doctor assigns to the patient’s expectation that the doctor will choose E. The exponent \( b \) qualifies the inequality-aversion coefficient into two cases: \( b = 0 \) when a person has not outcome-fairness motivation and \( b = 1 \) when she has. With the help of the equations (4) – (7) we can remodel the games G1 and G2 on the basis of the adjusted payoffs of the doctor’s non-compliant choices in table 3.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
\textbf{G1} & Patient &  \\
\hline
 & H & L \\
\hline
Doctor & E & 600, 600 \\
 & C & 700 - 1,6^b \lambda 100p, 500 \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
\textbf{G2} & Patient &  \\
\hline
 & H & L \\
\hline
Doctor & E & 600, 200 \\
 & C & 700 - 3,3^b \lambda 100q, 100 \\
\hline
\end{tabular}
\end{table}
As we can see, the parameters $\beta$, $b$, $\lambda$, $p$ and $q$ are in the center of our analysis. Parameters $p$ and $q$ represent second order empirical expectations. In the data analysis section we employ the distinction between *indirect* and *direct* second order empirical expectations in the meanings introduced in section 2.3. The relevant predictions of the model concerning the doctor’s choices can now be stated as follows:

P1) a majority of the doctors conform to the social norm by choosing E rather than C;

P2) in both games a higher *indirect second order empirical expectation* (ISE) implies a higher probability of the expensive treatment chosen by the doctor;

P3) in both games a higher *direct second order empirical expectation* (DSE) implies a higher probability of the expensive treatment chosen by the doctor;

P4) a portion of doctors express outcome fairness motivation and shift from the choice of C in G1 to the choice of E in G2.

In addition to these implications we comment on two additional hypotheses that can be derived from our theoretical framework. *First*, as suggested in section 2.1., some people may be more motivated to overthrow the expectations related to the salient social norm. Alternatively, they may be motivated to comply with some other social norm they see as salient or worthwhile to comply with. *Second*, we suggested in section 2.3. that direct second order expectations are sensitive to the opaqueness of the patient’s inner state. Both hypotheses are potentially worthwhile for closer exploration in further experiments.

5. Results of a doctor-patient game

5.1. Distributions in DPD game and DPT games

In DPD game each subject made, in the role of a doctor, a decision to allocate 1000 points between herself and the recipient (labeled as ‘patient’). The game is comparable to the general version of the dictator game where the average allocation to
the recipient has been found to vary considerably as a response to the framing of the situation. A context-free frame produces usually a two-peaked distribution with peaks at the minimum (or near it) and around the middle point. Usually the spread is from zero to the middle-point. A frame with a context makes the distribution towards one-peaked with the peak at the middle or even at the maximum (or near it). In the first stage of dictator games in the 1990s, when the basic robustness tests were conducted, it was found that the mean allocation (% of the whole pie) is 26.4%, varying between 10% to 50% (Camerer 2003). In more recent studies the range has expanded to the spread between 2.7% and 82.3% (Brañas-Garza 2006, 319).

Figure 3. Allocations in the DPD game

Figure 3 shows the distribution of the allocations in DPD game. We can see that the context of medical treatment has changed the distribution when set against a typical distribution in a context-free dictator game. The doctor-patient context is associated with a higher mean allocation percentage, higher share of those who gave about a half of the endowment, and considerable frequencies in allocations greater than the half. Thus, people in the doctor-patient context adopt a positive other-regarding behavioral pattern more frequently and allocate a greater sum of money to the recipient than in ‘abstract contexts’.

Table 4 shows the treatment choices by the doctors in games G1 and G2 in the DPT game. The 'selfish' choice of C predicted by the traditional non-other regarding

---

10 The percentages are based on 22 different experimental sessions in nine reports in Camerer (2003, table 2.4). Andreoni et al. (2007) gives a similar (25%) percentage to mean allocation.
preferences model is chosen only about 36% (29%) of the subjects in G1 (G2). We can conclude that the observed data are different from the traditional model’s prediction when the test proportion is 0.5179 indicating that the majority, that is, 51.8% of the subjects would choose C (sig. = .011 in G1 and sig. < 0.00 in G2). The observed distribution can be explained by a model with other-regarding preferences (see Linnosmaa et al, 2009) and by our model introduced in this paper. In general, we conclude that the data are in line with the prediction P1.

**Table 4.** Treatment choices by doctors in DPT game

<table>
<thead>
<tr>
<th></th>
<th>G1</th>
<th></th>
<th>G2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Abs.</td>
<td>%</td>
<td>Abs.</td>
<td>%</td>
</tr>
<tr>
<td>E</td>
<td>36</td>
<td>64.3</td>
<td>40</td>
<td>71.4</td>
</tr>
<tr>
<td>C</td>
<td>20</td>
<td>35.7</td>
<td>16</td>
<td>28.6</td>
</tr>
<tr>
<td>Total</td>
<td>56</td>
<td>100</td>
<td>56</td>
<td>100</td>
</tr>
</tbody>
</table>

**5.2. Influence of expectations and outcome fairness**

Regarding the predictions P2 and P3 we analysed, how norm-related expectations influenced on the choices the subjects made. The subjects were asked to assess (i) what percentage of the real world doctors follow the social norm (ISE) and (ii) the probability the patient assigned to the possibility that the doctor in the game chooses the expensive treatment (DSE_G1 and DSE_G2). As explained in sections 2.3. and 4.2. the former variable measures indirect second order empirical expectations (i-expectations, for short) and the latter two variables measure direct second order empirical expectations (d-expectations, for short).

We used a Mann-Whitney test to study whether those who chose E had significantly higher scores in ISE and DSE than those who chose C. The results are as expected: E-choosers have higher i-expectations both in G1 (z = -2.466, p = .006, 1-sided, exact test) and in G2 (z = -3.524, p = .000, 1-sided, exact test). Likewise, E-choosers have higher d-expectations both in G1 (z = -1.788, p = .037, 1-sided, exact test) and in G2.
(z = -2.008, p = .022, 1-sided, exact test). Also a point biserial correlation was undertaken in order to examine the strength of association between the treatment choices and expectations. The results in table 5 show that the i-expectations have a statistically significant moderate effect on treatment choices, and that the d-expectations have a statistically significant but small effect on choices.

Table 5. Test of point biserial correlation between treatment choices and expectations

<table>
<thead>
<tr>
<th></th>
<th>ISE</th>
<th>DSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>p*</td>
</tr>
<tr>
<td>G1</td>
<td>.328</td>
<td>.007</td>
</tr>
<tr>
<td>G2</td>
<td>.566</td>
<td>.000</td>
</tr>
</tbody>
</table>

*1-sided test

The prediction P4 suggests that a portion of doctors shifts from C to E when moving from G1 to G2. The related theoretical interpretation is that if a doctor has outcome-fairness motivation, the greater payoff-difference in G2 strengthens the doctor’s disutility from the deviation from the norm. Therefore she more probably chooses an expensive treatment in G2.

Table 6. Treatment choices in a ‘before-after’ table

<table>
<thead>
<tr>
<th></th>
<th>G2 (‘after’)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>G1 (‘before’)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>4</td>
<td>32</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>12</td>
<td>8</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>40</td>
<td>56</td>
<td></td>
</tr>
</tbody>
</table>

We can see from table 6 that 8 out of 20 (40%) doctors who chose C in G1 changed to E in G2. As a treatment effect this is not, however, statistically significant shift (McNemar exact test, sig. = 0.194).

From a different perspective, we studied whether the choices in G2 could be explained in terms of the ‘equal-split-orientation’ the subjects revealed in the modified dictator game we call the DPD game. To test this idea, we constructed a
variable (DISEP) which measures a subject’s distance $d$ from the equal division of 1000 points by calculating $d = |500 - r|$ where $r$ is the points the subject allocated to the responder.

A Mann-Whitney test was undertaken to examine whether those who chose E had significantly higher scores in DISEP than those who chose C. The result indicates that there is a difference in the case of G2 ($z = -2.327$, $p = 0.01$, exact test, 1-sided) but not in the case of G1 ($z = -1.035$, $p = 0.153$, exact test, 1-sided). A point biserial correlation was carried out to assess the strength of association between the treatment choices and the subject’s distance from the outcome-fairness defined as even division in the dictator game. The results shown in table 7 suggest that the relationship is moderate and statistically significant in the case of G2. This gives empirical support to P4.

Table 7. Test of point biserial correlation

<table>
<thead>
<tr>
<th></th>
<th>Distance from outcome-fairness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
</tr>
<tr>
<td>G1</td>
<td>-0.203</td>
</tr>
<tr>
<td>G2</td>
<td>-0.365</td>
</tr>
</tbody>
</table>

5.3. Comments on the deviations from expectations and on the sensitivity of DSE to opaqueness

Finally, few words are to be said regarding the choices of C in the DPT games and the sensitivity of DSE to the opaqueness of the inner state of the patient in the version G2 of the DPT game. As stated, C is a dominant strategy in the traditional game-theoretic reasoning with a non-other-regarding preferences model. However, the choices of C can quite easily be made predictable by a heterogeneous social preferences model such as Charness and Rabin (2002), for example.
Our discussion of the theoretical framework of the model of normative expectations in section 2 suggests two major motivations a person could have for deviating from a particular social norm. One is subversive tendency (see Hargreaves Heap and Varoufakis 2004) which would mean that some players in the game decided to frustrate the opposite player’s expectations or to act contrary to the salient norm. Another motivation would be related to another norm with other expectations: maybe some players thought that in such situations as the experiment was, it is a regularity to choose C. Figure 4 shows that among those who chose C, there are players whose $p_s$ was quite high ($\geq 0.50$). This would suggest that there were some subjects who decided to frustrate their patient’s expectations and comply with some rivaling social norm.

Regarding the sensitivity of DSE to the opaqueness of the inner state of the patient in the DPT game we see in table 8 that the mean expectation was higher in G1 (49.8%) than in G2 (43.7). In addition, we see that 21 subjects stated lower probability in G2 than in G1 as was suggested by our sensitivity hypothesis. A Wilcoxon signed-ranks indicates, however, that there is only a statistically symptomatic difference ($Z = -1.485, p = .070$, 1-sided exact test) between G1 and G2.
Table 8. Comparison of DSE in G1 vs. G2

<table>
<thead>
<tr>
<th>Within-subject reference values</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSE in G2 &lt; DSE in G1</td>
<td>21</td>
</tr>
<tr>
<td>DSE in G2 &gt; DSE in G1</td>
<td>13</td>
</tr>
<tr>
<td>DSE in G2 = DSE in G1</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>DSE in G1</td>
</tr>
<tr>
<td></td>
<td>DSE in G2</td>
</tr>
</tbody>
</table>

Albeit the sensitivity of DSE to opaqueness is only symptomatic in the statistical sense, the results give some additional information in interpreting our theoretical model. Namely, a lower level of direct second order empirical expectations in G2 than in G1 were associated to an increased frequency of norm-compliance in G2 than in G1. This may be interpreted either by referring to the increased payoff-difference in G2 compared to G1 or to a stronger influence of indirect second order expectations in G2 than in G1. This would mean that either (i) an outcome-fairness becomes ‘operative’ in G2 or (ii) the population-based expectations become stronger in G2, or that (iii) both factors are in operation.

6. Discussion

In this paper we have investigated a doctor-patient experiment and corresponding games between a doctor and a patient. As suggested, the setting can be generalized so as to serve as a stylized account of a specialist-client interaction. The main idea was that such interaction often involves quite strong and salient professional standards (regularities) and that one’s deviation from these standards gives rise to resentment which, in turn, explains generally why many people often find it profitable to comply with the standards.

The theoretical model introduced in section 3 allows that a decision-maker’s utility depends positively on her own monetary outcomes and negatively on the disapproval the others would target on her in the case she frustrates other's expectations. In addition, the structural payoff-difference influences on the strength of the possible dis-utility that enters into the total utility: the larger (smaller) the advantageous (disadvantageous) difference, the stronger impact the dis-utility will have. The model
combines some recent ideas related to normative expectations and inequality-aversion presented in separate models by Sugden (2000) and Fehr and Schmidt (1999) and to quilt-aversion model by Charness and Dufwenberg (2006) and Dufwenberg and Battigalli (2007).

The major finding of the empirical appraisal of the model is, that it explains, at least partially, how social norms (understood as regularities and/or normal practices) influence on subjects' behaviour in the experiment where the framing of the situation (including the labels of the roles of the players) highlights a particular norm. The subjects made, in the roles of a doctor, choices which are predicted by our model but not by traditional game-theoretic reasoning. In particular, our data supports the view that second order empirical expectations have a role in the subjects' behavior. If a subject believes that her opposite player expects her to choose the norm-compliant choice, it is also likely that the subject makes the expected choice. Also the distinction between two genres of second order expectations – those based on population distribution and the others based on one’s model of the opposite side’s internal state – was given some sidelight. Specially, the latter genre is sensitive to the opaqueness of the opposite side’s internal state as implied by the theoretical definition of the genre. These findings contribute to the recent endeavour by both theorists and experimentalists in economics to study the influences of social norms to better understand human economic behaviour.

In general, the above finding coincides with the findings by Bicchieri and Xiao (2009) and Faillo et al. (2008): the choices of the subjects are in line with their expectations. In a sense, then, there are answers from differing perspectives to the idea that some important outcomes of experiments which have been interpreted as resulting from other-regarding preferences, may be, indeed, results of norm-compliant behavior (see List 2007, Guala and Mittone 2009). However, there are some differences among these studies in the methodology of bringing norms and expectations into the experiment which make the more detailed comparison of the results fairly laborious. In short, Bicchieri and Xiao gave subjects information on other subjects’ choices in earlier similar experiments, and Faillo et al. gave subjects alternative principles and
corresponding rules among which the subjects were able to seek agreement by a voting procedure. In our experiment the framing of the game (by labeling and roles) was intended to highlight a social norm and to make it salient among the subjects.

In addition, another main finding was that the subjects' response to the experimental treatment-effect was towards behaviour predicted by the model: about 40% of the subjects who chose inexpensive in G1 shifted to expensive in G2. This can be interpreted either as an ‘activation’ of an outcome-fairness motivation or as an intensification of the indirect second order expectations in G2 when compared to those factors in G1. Naturally, both of these factors may have been in operation. One future challenge, therefore, is the identifying of these factors and the estimation of their significance in various context of decision-making.

7. Conclusions

In this paper we have focused on the influence of social norms and norm-based expectations on economic behavior. We have found that a formulation of an emotional sanctioning model explains the behavior of subjects in the expert-client interaction we called a doctor-patient game. In particular, the empirical expectations are in line with the decisions the subjects make when behaving in the role of a doctor in the game. In addition, we found that (i) the way in which the expectations are formed as well as (ii) the activation of the outcome-fairness motivation have influences upon decisions.

Like Bicchieri and Xiao (2009) we also want to bring forth the idea that empirical expectations have important implications for policy-making. In particular, rather than highlighting the detrimental consequences of unwanted behaviour the policy makers should emphasize that many, and often almost all others, comply with norm-based expectations. This idea could be applied in many domains in society's efforts to improve citizen's health behaviour or in policies of crime-preventing, for example. From of bit different perspective and on the basis of slightly different experiments,
Gächter (2007) speaks of belief management. Accordingly, also the way in which tax authorities, for example, publicly deal with tax evasion or tax reforms, or how top managers or politicians act, have an effect on the majority's view concerning what is regularly done in a particular situation.

References


