Responsibility and effort in an experimental labor market

Gary Charness

Department of Economics and Business, Universitat Pompeu Fabra,
Ramon Trias Fargas 25-27, 08005 Barcelona, Spain

Received 16 September 1998; received in revised form 2 August 1999; accepted 30 August 1999

Abstract

Previous indirect evidence suggests that impulses towards pro-social behavior are diminished when an external authority is responsible for an outcome. The responsibility-alleviation effect states that a shift of responsibility to an external authority dampens internal impulses towards honesty, loyalty, or generosity. In a gift-exchange experiment, we find that subjects respond with more generosity (higher effort) when wages are determined by a random process than when assigned by a third party, indicating that even a slight shift in perceived responsibility for the final payoffs can change behavior. Responsibility-alleviation can be a factor in economic environments featuring substantial personal interaction. © 2000 Elsevier Science B.V. All rights reserved.

JEL classification: A13; C91; D63; J20; J52

Keywords: Responsibility; Social behavior; Experiment; Generosity

1. Introduction

The responsibility-alleviation effect states that shifting responsibility for an outcome to an external authority dampens internal impulses toward honesty, loyalty, or generosity. Efficiency and performance may consequently be adversely affected. In this paper, evidence of this effect is reported in the context of an experimental labor market.

It is well-known that context and framing can affect decisions. This is particularly true when a decision is sensitive to social considerations and where norms for appropriate behavior are not clearly defined. One example is the dictator game, where one person in a mutually anonymous pair unilaterally chooses a binding allocation of a sum of money between the
two members of the pair. Here the division selected is notoriously sensitive to such issues as the phrasing of the instructions and the sense of entitlement. Another example is the Kahneman et al. (1986) survey, which examines pricing and labor compensation policies and demonstrates that one’s view of an action’s fairness depends on the context in which the action is chosen.

The intervention of an external authority can alter the context, effectively emphasizing one’s self-interest rather than issues of social responsibility. Griffith (1993) and Barkema (1992) find that performance is better without close supervision or monitoring. Campbell (1935), May and Loyd (1993) and Haines et al. (1986) find that an honor system induces more honesty than does a proctor.

It is certainly true that not every individual can be persuaded to assume responsibility for results or determinations. Although there is sometimes a tendency to avoid responsibility and commitment, it is also true that many people do agree to take on the responsibility for an unpleasant task. One may have a sense of ‘duty’ which acts as a motivation. A failure to heed this call may affect one’s self-reputation, with a concomitant loss of personal utility. Kerr and MacCoun (1985) provide evidence that people will even consciously let others free ride on their efforts if they feel that their social role prescribes it. This behavior happens even in the ant world: Rissing et al. (1989) find that ant queens who co-found a colony will accept an individually dangerous but socially-beneficial task (foraging for leaves) when directed to do so by an endogenous coordinating mechanism.

Social norms and fairness are probably most salient in environments featuring a high degree of interpersonal interaction (e.g. employment relationships, bargaining, and dispute resolution), where perceptions of what constitutes appropriate behavior may well affect individual choices. For example, Akerlof and Yellen (1986) claim that norms of fairness have significant effect on employment and wages in seemingly competitive labor markets. In their view, workers have a belief about what constitutes a fair wage. An employee would be expected to provide less effort when the actual wage is less than the perceived fair wage. The Akerlof (1982) gift-exchange model characterizes the offer of employment as an offer to ‘exchange gifts,’ where the worker’s effort level indicates the size of the reciprocal gift.

Gift-exchange experiments, as suggested by Akerlof (1982), are generally framed as non-repeated interactions between subjects in the roles of employer and employee and offer insight into non-strategic generosity and reciprocity. An employee receives a wage and then chooses an effort level. Any effort level above the minimum is costly to the employee but greatly benefits the person in the role of employer. Fehr et al. (1993) and Fehr et al. (1998) demonstrate a strongly significant positive slope for the effort–wage relationship in contrast to the predictions of the standard economic model. However, these studies only consider the case when an employer has actually chosen the wage.

Charness (1996) finds that the slope of the effort–wage relation is significantly steeper when the wage is chosen by the employer than when the wage is determined by an external process (either by a draw from a bingo cage or by the experimenter). Yet the slope is still

---

2 Hoffman et al. (1996) offer some nice illustrations of the sensitivity of choices in the dictator game.

3 They state (p. 261) that “if people do not get what they think they deserve, they get angry.”
quite positive with an external wage-determination process, suggesting that behavior is partially driven by a gain-sharing motive or generosity.

According to the principle of responsibility alleviation, we should expect impulses toward generosity to be diminished if a third party has assigned the wage, as an employee may perceive it to be a personally-sanctioned entitlement rather than an accident. Although it is true that the responsibility for effort provision remains with the employee, the initial wage choice defines the set of possible payoffs. The human intervention may enable the employee to rationalize shifting some of the responsibility for the final outcome onto the shoulders of the third party. Conversely, a subject in the random treatment cannot avoid accepting full human responsibility for the final allocations.

The experimental results reported below indicate a small but significant increase in effort provision when the wage is determined by a draw from a bingo cage. As it is difficult to credibly shift responsibility in an artificial laboratory environment (and here responsibility is only partially alleviated), the difference in effort levels we observe may represent a lower bound on the magnitude of this effect. In any case, the level of costly effort provision is lower when a neutral human agent chooses the employee’s wage and high levels of effort are much rarer. The resulting productivity is higher when an employee cannot attribute any responsibility for his wage to a neutral third-party and is the only human influence on the ultimate allocation of payoffs between employer and employee.

The social welfare effects of responsibility alleviation are greater when one considers dynamic outcomes, rather than one-shot static results. In a workplace, pro-social behavior can lead to improved relationships and increased productivity; in the educational realm, less cheating can lead to a greater respect for the learning process and increased scholarship; diminished adversarial behavior in disputes can effect a decrease over time in the number of disputes.

2. Previous theory and evidence

There is no previous general formulation of responsibility alleviation but we can find many examples of behavior that support the premise. One related idea is the distinction between extrinsic and intrinsic motivation. Deci and Ryan (1985) and Lepper et al. (1973) find that if people are given external rewards for pursuing an activity, their existing internal motivation is undermined. A relevant area is job performance, where there is empirical evidence that people may resent the presence of a monitor and that job performance may be adversely affected. Griffith (1993) tested monitoring and job performance, finding that performance was lower with physical monitoring than with no monitoring, except when a supervisor is actively monitoring.4 Barkema provides evidence from managers in Dutch firms which suggests that executive performance is better without close supervision. Baker et al. (1988) stress the role of trust and loyalty.

4 This is not meant to imply that simply reducing monitoring will inspire loyalty. The excellent study by Nagin et al. (1999) shows that telephone-sales employees provide less effort when there is less monitoring. However, the employee always expects substantial monitoring and so the “game” remains the same: the company does not really trust you, so it is ok to get away with what you can.
Some theoretical studies may help explain these results. Akerlof and Yellen (1986) argue that workers may acquire sentiments towards the firm by which they are employed, due to which their behavior can be affected. Frey (1993) provides a model which illustrates how external intervention, in the form of high levels of monitoring, can negatively impact work performance when personal relationships are important. Principal-agent theory presumes that the disciplining effect of monitoring has a beneficial effect on performance. However, an implicit psychological contract may exist between principal and agent, causing the agent to view intensive monitoring as evidence of distrust. This perceived distrust reduces the agent’s marginal benefit from work effort and encourages more opportunistic behavior, as issues of self-reputation are alleviated and an agent’s intrinsic motivation is reduced. Monitoring ‘crowds out’ work effort, particularly when personal relationships are salient.

Evidence of a slightly different nature is found in studies of the honor system. A seminal work by Campbell compares the behavior of students under an honor system and a proctor system. He concludes (p. 72) that “The amount of classroom dishonesty... was greater among the students working under the proctor system than among similar subjects working under the honor system.” Students strongly favored the honor system and were also optimistic regarding the possibility of developing a high moral standard among college students. May and Loyd (1993) also support the conclusion that the existence of an honor code is associated with increased academic honesty. In a similar vein, students in the Haines et al. (1986) study felt that the honor system actually instilled honesty and integrity in students.

In the area of dispute resolution, arbitration studies (e.g. Lester, 1989; Currie and McConnell, 1991) indicate that placing the responsibility for a decision in the hands of a third party can adversely affect the negotiation process — the expectation of third-party intervention may lead to impasse more frequently than would otherwise occur. McEwen and Maiman (1984) find that defendants in small claims court in Maine were nearly twice as likely to comply fully with mediated outcomes, as with judgments imposed by the court after adjudication.

MacCoun et al. (1992) state that “legal procedures that rely on formal, adversary adjudication can damage ongoing relationships.” Methods which promote cooperation in the resolution process may generate less hostility between the parties, preserving and sometimes strengthening relationships. This has clear implications for the likelihood and character of continuation disputes between these parties. Many studies (e.g. Thibaut and Walker, 1975; Tyler, 1990) indicate that procedural satisfaction may be as important as outcome satisfac-

---

5 In this model, agents maximize utility by choosing optimal effort \( E^* \), given monitoring \( M \) and the resulting benefits \( B(E,M) \) and costs \( C(E,M) \). Differentiation of the optimality condition gives: \( \frac{dE^*}{dM} = \frac{BEM - CEM}{CEE - BEE} \). By principal-agent theory, the disciplining effect would mean that \( CEM < 0 \). The crowding-out effect would mean that \( BEM < 0 \). The marginal cost of effort increases with effort (\( CEE > 0 \)), as does the marginal benefit from effort (\( BEE < 0 \)). The sign of \( \frac{dE^*}{dM} \) thus depends on the relative values of these derivatives; where the crowding-out effect is strong and the disciplining effect is weak, tighter monitoring reduces the agent’s effort.

6 An overall cheating measure used in this article shows 23.7 percent of the students cheating under an honor system versus 54.1 percent cheating with no honor system. The authors also suggest that a necessary ingredient is the internalization of the values espoused in the honor system.

7 In mediation, participants help shape the resolution and must consent to it; in arbitration, an outcome is imposed by a third party who hears the evidence. The authors take pains to minimize the problem of self-selection.
tion. A satisfied party is more likely to maintain a positive, productive relationship with others in the environment, and to refrain from sabotage, physical violence and other costly forms of passive or active rebellion. If parties with stakes in the outcome also have considerable responsibility for determining the outcome, process satisfaction is enhanced and resolution is more effective.

Taken together, these studies from a variety of fields suggest that any inherent sense of ‘fair play’ may be circumscribed if one does not feel the responsibility for an outcome. In the gift-exchange experiment described below, there is a subtle difference in the degree of responsibility a responder has for the final allocation of payoffs and so we might expect a modest difference in the results.

3. Experimental method

The experiment was conducted at UC-Berkeley and involved a simulated labor market. All the subjects were students at UC-Berkeley. A full description of the instructions and record sheets is available upon request. Seventy-eight subjects participated in the treatments where the employer did not determine the wage; 39 had the role of ‘employee’ and 39 had the ‘employer’ designation. Average earnings (including a US$ 5 show-up fee) were between US$ 16 and US$ 17 for a session of about 100 min.

Participants initially assembled in one room and were randomly divided into groups of ‘employers’ and ‘employees’ and then separated into two rooms. Having all the participants meet in one room at the outset helped to ensure that another real person’s payoffs were credibly dependent on the employee’s action choice.

As in the Fehr et al. (1998) BGE/TC treatment, each employee was paired anonymously with one employer in each period (10 periods in all) and it was common knowledge that workers and firms were not re-matched with the same person. The only changes were that the wage was not chosen by the employer and that the employee could not reject the wage.

In two sessions, each employee in each period was given a wage that was assigned by a draw from a bingo cage. Balls were drawn individually, in front of successive subjects. These identical wages were also used in two subsequent sessions. However, in these sessions the wage for each employee in each period was read (privately) from a sheet given by the experimenter. The employees were told that these wages had been selected by a third party (the experimenter). Each employee was in only one of these two treatments.

In all the cases, once assigned a wage an employee was asked to record an effort choice (between 0.1 and 1.0, inclusive) on a record sheet and to fold this sheet before turning it in to the experimenter. This sheet was then given to his or her employer in the other room, so that the employer had salient physical evidence of the employee’s choice. In a certain

---

8 In general, there were 10 employers and 10 employees in each session.
9 I thank Ernst Fehr for the latter suggestion.
10 Wages chosen in the employer-volition case in Charness (1996) were mapped onto the numbers of the balls in the bingo cage. If a particular wage (rounded to the nearest multiple of five, with employer-chosen wages >90 mapped to 90) had been selected by employers 4 percent of the time, then three of the 75 balls would correspond to that wage. Wages increased with increasing ball numbers.
sense, these cases without employer volition are analogous to a dictator game, since one party has no choice of actions, but is completely at the mercy of chance and other people.

The combination of wage and effort determined outcomes and monetary payoffs for each pair of subjects in a period. Each employer was given an endowment of 120 ‘income coupons’ in each period. The employer’s payoff function was given by:

\[ \Pi_F = (120 - w)^e \]  

(1)

where \( e \) denotes the employee’s effort and \( w \) the wage. The payoff function for an employee was defined as:

\[ \Pi_E = w - c(e) - 20 \]  

(2)

where \( c(e) \) is the cost of effort, a function increasing in \( e \). The minimum effort level of 0.1 had zero cost. These payoff functions and the values of the parameters were common knowledge and participants were required to calculate both employer and employee payoffs in three exercises with hypothetical wage-effort pairs. These exercises were reviewed before proceeding with the experiment, ensuring that subjects understood the payoff mechanism and that higher employee effort meant higher employer earnings but lower employee earnings. At the conclusion of a session, all the participants were paid privately.11

4. Results

If the perception that a neutral person (rather than a random mechanism) determines the wage causes an employee to feel less responsible for the employer’s payoff, the principle of responsibility-alleviation predicts that employees will provide higher effort levels when a wage has been randomly determined than when this same wage is determined by a human third party. This hypothesis can be tested statistically. As the assigned wages in the two treatments are nearly the same (the mean wage was 57.0 in the random treatment and 57.9 in the third-party case), we can compare effort levels with the Chi-square test (Table 1).

The test statistic is \( \chi^2 = 20.6 \) (d.f. = 9), so that the null hypothesis of no difference can be rejected at \( p = 0.02 \). We can also reduce the number of effort categories to low (0.1, 0.2, or 0.3), medium (0.4, 0.5, 0.6, or 0.7) or high (0.8, 0.9, 1.0) as in Table 2. The full results are shown in Table 3.

Here we have \( \chi^2 = 16.3 \) (d.f. = 2), so that the null hypothesis of no difference can be rejected at \( p = 0.01 \). The most dramatic difference between the treatments is that high effort provision is far more common with random wage determination — the test of the equality of proportions (normal approximation to the binomial distribution — see Glasnapp and Poggio, 1985) finds this difference is significant at \( p < 0.001 \) (\( Z = 3.65 \)).

The experimental design permits direct point-by-point comparison of the effort choices made by employees in the random and third party treatments, since the wages assigned to an employee in one treatment were identical (in sequence, as well) to those assigned to a corresponding employee in the other. If the wage-generating mechanism has no substantial

11 Experimental ‘guilders’ were converted to dollars at the rate of 25 guilders = US$ 1.
Table 1

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Effort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.1</td>
</tr>
<tr>
<td>Random</td>
<td>70</td>
</tr>
<tr>
<td>Third-party</td>
<td>77</td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Effort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>Random</td>
<td>114</td>
</tr>
<tr>
<td>Third-party</td>
<td>115</td>
</tr>
</tbody>
</table>

effect on the effort choice, we should expect the number of instances where the effort level selected in the random case exceeds that chosen in the corresponding third party case to be approximately equal to the number of instances where the reverse is observed. Let $N_r$ be the number of observational pairs where the effort chosen in the random case exceeds that chosen in the third party case and $N_t$ be the number of cases where the reverse is true. The data show that $N_r = 82$ and $N_t = 62$, while the chosen effort level was identical in 36 cases. 12 The normal approximation to the binomial distribution finds $Z = 1.67$, with $p < 0.05$ (one-tailed test). 13

Thus, it appears that people feel less of an impulse to contribute to the welfare of an anonymous employer when a third party is perceived to have in some way approved the wage, shifting some of the responsibility for determining final outcomes. While this is a somewhat subtle effect, the occasional impulses toward high levels of generosity (effort) seen in the bingo-cage treatment all but disappear when a third party has assigned the wage.

5. Conclusions

We have seen that the issue of responsibility can be an important determinant in an individual’s choice of actions. Third-party intervention appears to have a slight but definite, negative effect on the effort levels chosen by employees. This effect is not induced by strategic expectations of a party’s presumed ability to affect the third party decision, 14 but reflects a difference in employee generosity, perhaps induced by a disparity in relative

12 Although in principle there should be 190 joint observations of effort levels chosen (19 employees × 10 periods), due to some production errors there are only 180 precise matches.

13 The mean effort levels for each treatment were 0.3463 and 0.3220 for the random and third-party cases, respectively. As average wage differs slightly across conditions (higher average wages in the random case), a better measure is the ratio of discretionary effort (effort — 0.1) to discretionary wage (wage — 20). Overall, this ratio is 14 percent higher in the random treatment (0.00666–0.00586).

14 This contrasts with the ‘chilling effect’ (parties’ expressed positions diverge, hindering settlement) in arbitration, which is induced by beliefs that the third party will tend to ‘split-the-difference’ between positions.
Table 3
Wage/effort pairs

<table>
<thead>
<tr>
<th>Wage</th>
<th>Effort</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>1</td>
</tr>
<tr>
<td>0.6</td>
<td>1</td>
</tr>
</tbody>
</table>

Random-generated wages

20  17  3  2
25  2  1  1  1
30  2  1
35  2  2
40  9  1  3  2
45  5  2  1
50  4  1  2  3  1  2
55  1  1  1
60  10  4  7  10  5  1  1  3
65  1  2  1
70  8  4  1  2  7  4  1  1
75  4  3  1  2  3  2
80  4  1  4  3  1  1  4  1  1
85  5  2  1  1
90  2  1  2  1

Third-party-generated wages

20  14  1
25  3  2
30  1  2
35  7  1  1
40  9  3  4  1
45  1  2  3  2
50  3  4  5  2
55  2  1
60  13  2  4  7  14  2  1
65  2  1  1
70  10  2  2  5  10  2  2
75  2  2  1  1  2  3
80  10  1  2  2  7  1
85  1  1  1  1  1
90  1  1  2  1

payoffs. As in Kahneman et al. (1986), an identical action (a wage) produces differing responses (effort levels) depending on the context in which it is viewed. This context may affect perceptions of applicable social norms and beliefs about appropriate behavior. Similarly, self-regulation and personal participation may lead to more socially-beneficial outcomes in other environments.

The experimental results indicate that, in general, one is more generous with an anonymous stranger when one must assume full human responsibility for an allocation of payoffs. When an external and neutral party can be seen to have at least some responsibility for the final outcome, the parties directly involved tend to be less concerned about the well-being of others. Third-party intervention may be helpful in many situations, but if people are less
generous or more partisan when a third party is perceived to have some responsibility for allocations and payoffs, there can be a high social cost from such intervention.

The responsibility-alleviation effect can have major consequences for the design of employment environments and dispute resolution processes. As relationships are important in the labor context, enhancing cooperative behavior and loyalty is clearly beneficial. While benefits of enhanced pro-social behavior can be substantial in the static case, the potential impact potential is greater in a dynamic context, particularly in economic environments featuring repeated personal interaction. There are implications for management in the workplace, for the design of dispute resolution systems, and for educational testing policies. Further experimental research would help delineate the scope of this effect.

Acknowledgements

Financial support from the Vice-Chancellor for Research Fund at UC-Berkeley is gratefully acknowledged. I thank George Akerlof, David. I. Levine, Rob MacCoun, Ted O’Donoghue, Matthew Rabin, and an anonymous referee for helpful comments.

References