Peer Pressure in Work Teams:
The effects of Inequity Aversion

David Masclet
Décembre 2002
Peer Pressure in Work Teams:  
The effects of Inequity Aversion

Pression des pairs dans les équipes de travail:  
Les effets de l’aversion à l’inégalité

David Masclet¹

Abstract:
Many empirical studies have shed light on the efficiency of peer pressure. I propose here to model peer pressure by incorporating in the utility function self centered inequity aversion. I find that opportunity for sufficiently inequity averse players to punish their peers, is effective in inducing others to cooperate. At the equilibrium, all players cooperate and punish any shirker since punishing is a way to reduce inequity. Contrary, nobody cooperates without peer pressure even if players are inequity averse.

Key words: Peer Pressure, Work Teams, Inequity Aversion, Fairness, Cooperation

JEL : A13, C72, D63, D20, L23

Résumé:

Mots Clés : Pression des pairs, équipes de travail, Aversion à l’inégalité, bienveillance, coopération

JEL : A13, C72, D63, D20, L23

¹Masclet David : Université Lumière Lyon, GATE, masclet@gate.cnrs.fr. I thank Marie-Claire Villeval, Laurent Flochel, Irene Valsecchi and the participants at the French Economic Association Conference, Lyon for their helpful comments. I also thank Nathalie Colombier and Stephane Mahuteau for their precious help. Remaining mistakes and eccentricities are only fruit of my own incompetence.
1. Introduction

Many empirical studies have shed light on the efficiency of peer pressure. Peer pressure can be defined as mechanisms of mutual monitoring and sanction established within a group of agents by the agents themselves in order to dissuade the members of the group to adopt a non-co-operative behavior. For example, authorities often post reward for citizens who report violations of crimes. Similarly, monitoring by peers in work teams, credit associations (Stiglitz, 1990; Varian, 1990), partnerships (Jones and Svejnar, 1985; Rehder, 1990), is an effective means of attenuating incentives problems that arise when individual actions affecting the well being of others are not subject to enforceable contracts.

Experimental studies also provided additional supports for the positive effect of peer pressure in inducing high levels of contribution to a public good. In their paper, Fehr and Gaechter (2000) consider how the level of contribution to a public good responds to the availability of an option for individual agents to punish other agents after they have observed their decisions. The results of Fehr and Gaechter are striking. In the only subgame perfect Nash equilibrium of the game it cannot be a best response to punish in the second game because such punishment is costly to the punisher. By backward induction, the best response in the first stage is to contribute zero to the group account. However, they observed a willingness to punish on the part of subjects. A possible explanation for such punishing behavior in one-shot conditions is that subjects would be willing to sanction for unfair behavior.

In recent years experimental economists have gathered evidence that fairness cannot be ignored in social interactions (Guth, Schmittberger and Schwarze, 1982). One can distinguish two main approaches of fairness. The first approach relies on the importance of intentions of the other agents as a source of reciprocal behavior (Rabin, 1993; Dufwenberg and Kirchsteiger, 1998). These models provide several interesting insights, but are not well suited for predictive purposes (Fehr and Schmidt, 2001). Moreover, the multiplicity of equilibria is a general feature of these models, which makes the equilibrium analysis fairly complex even in extremely simple sequential games. The second approach of fairness assumes that agents have social preferences; i.e. the agents’ utility function does not only depend on their own material payoff but also on how much the other agents receive. Fehr and Schmidt (1999) develop a model in which a person exhibits inequity aversion if he dislikes being better and/or worse than others. At the same time, Bolton and Ockenfels (2000) propose a similar model of inequity aversion in which people compare their material payoff to the material average payoff of the group. Even if models of social preferences are not fully satisfactory since they assume that agents are only concerned about the distributional consequences of their acts but not about the intentions, one can probably go a long way with such very simple models.

The aim of this paper is to propose a model of peer pressure which incorporates fairness in the utility function. I consider a work team of two agents. In a first stage, the

---

2 The fundamental difference between Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) is the following : BO assumes that subjects like being as close as possible to the average payoff while FS assumes that subjects dislikes a payoff difference to any other individual. Engelmann and Strobel, (2000) designed a very simple experiment in order to compare the relative performance of the competing theories by BO and FS. Their results were clearly in favor of FS.
agents decide to cooperate or to defect. In a second stage, they observe the decisions of their peers and can decide whether to punish them or not. Following Fehr and Schmidt, 1999, I assume that people are motivated by self-centered inequity aversion. The agents’ utility depends not only on the payoffs of the original standard game but also on inequity aversion. An individual is inequity averse if he incurs disutility both from being worse off in material terms than the others (disadvantageous inequity) and from being better off (advantageous inequity). I show that the threat to punish defectors may be credible if players are sufficiently upset by inequality. Then it induces potential defectors to cooperate.

This analysis builds on the work of Kandel and Lazear (1992) with the notable exception that I adopt here a sequential game-theoretic approach of peer pressure. Thus, Barron and Gjerde (1997) is much closer to this work in that they also adopt a sequential approach. However, their sequential game interpretation of peer pressure assuming that agents choose work effort taking as given the peer pressure environment, is not completely satisfactory.

The remaining of the paper is organized as follows. In Section 2, I present a model of inequity aversion without peer pressure. I introduce peer pressure in section 3 and analyze the impact of inequity aversion on cooperation. Section 4 discusses the assumptions. Section 5 concludes.

2. Model without peer pressure

2.1. Assumptions and payoffs

At the beginning of the game, players form a group to produce a good. If they Cooperate, more of the good is produced than if they defect. Defection may be thought of as shirking. Let \( n \geq 2 \) denote the number of members in the group and \( m, m \leq n \), the number of members who cooperate. All output is split equally among the members. The average output per member is denoted by \( f(m,n) \) and the cost to a single member of cooperating is denoted by \( X \). When all members cooperate, each member receives a payoff of \( f(n,n)-X \). If some members defect, the payoff to each defector is \( f(m,n) \) and the payoff to each cooperator is \( f(m,n)-X \). When all defect, each member receives a payoff of zero. The payoff to each member \( i \) is therefore:

\[
x_i = \begin{cases} 
  f(m,n) - X & \text{if } i \text{ cooperates} \\
  f(m,n) & \text{if } i \text{ defects} \\
  f(0,n) = 0 & \text{if all players defect}
\end{cases}
\]

(i)

I examine the equilibrium under the following key assumptions:
(1) **free riding problem.** A defector gets a higher payoff than a cooperator. Defecting is a dominant strategy so that for any number of cooperators \(m > 0\),
\[ f(m, n) - X < f(m - 1, n) \]
(ii)
I also require that
\[ f(m - 1, n - 1) < f(m, n) \]
(iii)
which is the mathematical statement that there exists aggregation economies: the presence of an additional cooperating member raises per capita output. Setting \(m = n\), the above inequity implies that per capita payoffs are larger in a larger cooperating group. The output is normalized with zero cooperators to \(f(0, n) = 0\), which satisfies “No aggregation economies without cooperation”. As a complement to the above inequity, I assume that the presence of a defecting member does not raise the average output. Therefore,
\[ f(m, n - 1) = f(m, n) \]
(iv)
(2) Let us assume that ‘cooperation’ is socially valuable, so that the per capita payoff net of costs is higher if all cooperate than if all defect,
\[ f(n, n) - X > 0 \]
(v)
(3) Finally, following Fehr and Schmidt (1997), I also assume here that people are motivated by self-centered inequity aversion. The disutility from inequity is self-centered in the sense that player \(i\) compares himself to each of the other players but he does not care per se about inequalities within the group of his opponents. I transform here a standard game into an «inequity aversion game». In this game the players ‘utility depends not only on the payoffs of the original standard game but also on inequity aversion. An individual is inequity averse if he incurs disutility both from being worse off in material terms than the others (disadvantageous inequity) and from being better off
(advantageous inequity). I consider that the players are more sensitive to disadvantageous inequity which is given by the inequity aversion term $\alpha_i$, $\alpha_i > 0$ than to advantageous inequity, given by $\beta_i$ with $0 \leq \beta_i < 1$ such that $\alpha_i > \beta_i$. The utility function of player $i \in \{1, \ldots, n\}$ is then given by:

$$U(x_i) = x_i - \alpha_i \frac{1}{n-1} \sum_{j \neq i} \max(x_j - x_i, 0) - \beta_i \frac{1}{n-1} \sum_{j \neq i} \max(x_i - x_j, 0)$$

(1)

The first term of equation (1) represents the pecuniary payoff of player $i$. The second term in (1) measures the utility loss from disadvantageous inequity. The third term represents the utility loss from advantageous inequity. Furthermore, the assumption $\alpha_i > \beta_i$ captures the idea that a player suffers more from inequity that is to his disadvantage. For simplicity I assume that the utility function is linear in inequity aversion as well as in $x_i$. This implies that the marginal rate of substitution between monetary income and inequity is constant.

2.2. Equilibrium

In this section, I examine the symmetric Nash equilibrium in which all players choose the strategy of defection. Without opportunity of sanction, all players always defect is a Nash equilibrium. Indeed when the players have no opportunity to punish each other they are not able to support cooperation even if players are inequity averse. This follows from the fact that the only mean for an inequity averse player to reduce this inequity, is to reduce his own effort.

**Proposition 1**: without opportunity of sanction, there exist an equilibrium without cooperation and the output is not produced.$^3$

**Proof**: I prove that the strategy in which nobody cooperates supports an equilibrium. To do so, I must show that no subject has an incentive to deviate from this proposed equilibrium. Let us consider the utility function of player $i$ when all other players defect. The payoff of each subject $i$ can be written as:

$$U_i(X = 0) : f(0, n) = 0$$

(2)

Suppose that player $i$ deviates from the equilibrium and chooses to cooperate. Then his payoff is given by:

$$U_i(X > 0) : f(1, n) - X - \frac{\alpha_i}{n-1}(n-1)X$$

(3)

The first term of the utility function of player $i$ represents the average output per member player $i$ receives when all the other subjects defect while he cooperates. The second term represents the cost for player $i$ of cooperating. Finally, the third term represents the

---

$^3$ The formal proof for the uniqueness of the equilibrium without cooperation under weak conditions of inequity is relegated to appendix A.
disadvantageous inequity the player $i$ experiences relative to the $(n-1)$ other players who fully defect. Indeed, the net gain from defecting is $X$.

Let us consider the incentive for player $i$ to deviate. Player $i$ is incited to cooperate if and only if the following condition is verified:

$$U_i(X > 0) - U_i(X = 0) > 0 \iff \{ f(1,n) - X - \alpha X \} - \{ f(0,n) \} > 0$$

(4)

From assumption (ii), one knows that $f(1,n) - X < f(m-1,n)$. Thus it appears that $f(1,n) - X < f(0,n)$. Moreover, when subject $i$ cooperates while all the other subjects defect, he experiences disadvantageous inequity relative to all the $(n-1)$ other players. Since, $\alpha > 0$, then the above condition is never verified. Therefore, nobody is incited to deviate from the equilibrium. Thus cooperation cannot emerge in a team without peer pressure. Indeed, players are not incited to cooperate since they incur disadvantageous inequity relative to the other players who fully defect. The theoretical predictions remain the same as in case without aversion to inequity.

3. Model with peer pressure

Consider now a slightly different game in which the players have the opportunity to punish their peers. Can cooperation emerge when inequity averse players have the opportunity to sanction their peers?

3.1. Assumptions and payoffs

In a first stage, the players decide to cooperate or to defect. In a second stage, they observe the decisions of their peers and can decide whether to punish them or not. When a player is punished by someone, he receives a sanction $Y$. A player who punishes his peer also incurs a cost for doing that given by $cY$ with $0 < c \leq 1$.

(4) In addition to the general assumptions, I must restrict the magnitude of the parameters. Let us assume that :

$$Y_i > X$$

(vi)

so that the sanction must be high enough such that deviators cannot get a higher payoff than the enforcers. The penalty is larger than the benefit from cheating against cooperators.
3.2. Equilibrium

In this section I examine a symmetric perfect equilibrium in which each player adopts a strategy called “cooperation enforcement”. A player following strategy of “cooperation enforcement” cooperates along the equilibrium path and punishes anyone who deviates from this strategy unless he has violated the strategy of cooperation enforcement in the preceding phase, in which case he does not punish. On the equilibrium path, everyone cooperates; if anyone deviates by defecting, the others still cooperate but punish him. I prove that “cooperation enforcement” supports an equilibrium with cooperation. Indeed, contrary to the case without any opportunity of punishment, there is now another means to reduce disadvantageous disutility: players can interfere on the payoff of the others players by sanctioning them.

**Proposition 2**: The opportunity to reduce the payoffs of the other members increases cooperation if players are sufficiently upset by inequity.

**Proof**: Proposition 2 show that cooperation can be sustained as an equilibrium. I use backward induction to verify that the strategy combination in which all players cooperate to produce the output in first stage and punish any defector in second stage is a subgame perfect equilibrium.

(1) In the second stage, the players decide to punish or not the other members of their group. Let \( n \geq 2 \) denote the number of members in the group and \( m \), the number of members who cooperate. I have to check that the threat of punishment is credible. To do that, I must show that a subject is not incited to deviate from the subgame equilibrium in which subjects punish any defector. Let us consider the utility of a potential enforcer who does punish shirkers while all other enforcers punish shirkers such as:

\[
U_i(P > 0) = f(m,n) - cY(n-m) - \frac{\alpha_i}{n-1} \{(n-m)(X + cY(n-m) - mY)\}
\]

(5)

The first term of the utility function of player \( i \) represents the average output per member player \( i \) receives. The second term represents the cost for player \( i \) of cooperating. The third term represents the disadvantageous inequity the player \( i \) experiences relative to the \((n-m)\) players who fully defect. The sanction \( Y \) each enforcer imposes on each shirker reduces the disadvantageous inequity from \( Y \). However, subject \( i \) incurs a cost \( cY \) when he sanctions a shirker, which increases the disadvantageous inequity. At least, the last term represents the cost of punishing the \((n-m)\) shirkers.

---

\(^4\) Hirshleifer and Rasmusen, (1989)
Let us consider now the utility of player $i$ if he chooses not to punish a shirker such as:

$$U_i(P = 0) : f(m,n) = -\frac{\alpha}{n-1}(n-m)(X - (m-1)Y) - \frac{\beta}{n-1}(m-1)(n-m)cY$$

(6)

If player $i$ chooses not to punish the other members of his group, he saves the cost $cY(n-m)$ but experiences now more disadvantageous inequity relative to the defectors and also advantageous inequity relative to the $(m-1)$ other enforcers who fully punish the $(n-m)$ defectors. The latter generates a utility loss of $\left[\frac{\beta(m-1)(n-m)cY}{(n-1)}\right]$ whereas the former reduces utility by $\left[\frac{\alpha(n-m)Y(1-c(n-m))}{(n-1)}\right]$. Thus a player chooses to punish the defectors if and only if the following condition is verified:

$$\frac{1}{n-1}\left[\alpha(1-c(n-m)) + \beta(m-1)c\right] > c$$

(7)

Some simple algebraic manipulations show that the above condition is equivalent to the following:

$$g(c,m,n,\alpha,\beta) = \alpha(1-c(n-m)) + \beta(m-1)c - c(n-1) > 0$$

(8)

The following table resumes the expected effects of the different parameters of the equation (8) (Appendix B).

<table>
<thead>
<tr>
<th>$n$</th>
<th>$c$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>g(.)</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Thus if players are sufficiently averse to disadvantageous inequity but also to advantageous inequity and if the cost of punishing a shirker is sufficiently low, then players will punish all defectors.

(2) Suppose that the condition (7) is satisfied. In first stage of the game, the players decide to cooperate or not. Consider now the incentives of a player to deviate in first stage. Let us consider first the utility of a player who follows the equilibrium path, i.e. that he cooperates while all the other players cooperate. Then his payoff is given by:

$$U_i(X > 0) : f(n,n) = X$$

(9)

If player $i$ deviates from the equilibrium and defect at this stage, his payoff becomes:
\[ U_i(X = 0) : f(n-1,n) - \frac{\beta}{n-1}(n-1)X - (n-1)Y \]

(10)

Then \( i \)'s immediate payoff from defecting is the receipt of \( f(n-1,n) \) instead of \( f(n,n) - X \). But if he defects, he will be punished in second stage, which yields him a loss of \( (n-1)Y \). Therefore the condition for player \( i \) to prefer cooperate is:

\[
 f(n-1,n) - \frac{\beta}{n-1}X - (n-1)Y < f(n,n) - X
\]

(11)

Assumptions (iii) and (iv) together imply that \( f(m-1,n) < f(m,n) \) (per capita is raised by adding a cooperator) and (vi) \( X < (n-1)Y \), so condition (11) is always satisfied. Hence the players are not going to deviate at stage 1. Then, in the first stage each player cooperates. In the second stage, there is no punishment since each player cooperates in first stage. If one cooperators deviates and defects, then each enforcer punishes the defector. Comparing the cases with and without opportunity of punishment, it is easy to see that cooperation is greatly improved if there is an opportunity to punish defectors.

There is another symmetric perfect subgame equilibrium in this sequential game in which nobody cooperates at first stage and nobody punishes at second stage since it provides the same material payoff for each players and does not generate inequity. Which equilibrium will be chosen? Let us note E1, the equilibrium with cooperation and E2 the equilibrium without cooperation. E1 Pareto dominates E2. In a single shot interaction, it is difficult to say which equilibrium will be chosen. However, if one applies Pareto dominance without perfection, E1 will be the equilibrium since both players prefer it (Rasmusen, 1994).

Suppose now that this game is played twice with the first game outcome observed before the second game begins. A possible refinement consists in using the Pareto dominated equilibrium as a threat in case of non co-operation. Suppose further that the players anticipate that the second game outcome will be as follows: E1 if the first game outcome is E1; E2 if the first outcome game is E2 (Gibbons, 1992, pp82-88). Then E1 is a subgame perfect equilibrium for both games of this super game. Since only Nash equilibrium are discussed, the players’ agreements are self enforcing which is a more limited suggestion than the approach in cooperative game theory according to which the players make binding agreements\(^5\).

\(^5\) A generalized proof of this result, which can also apply to the repeated plays, is given in van Damme, (1991, pp. 198-206), in Gibbons, (1992, pp. 82-88) and in Eichberger (1993, pp. 224-229).
Thus under some conditions, cooperation can be sustained as an equilibrium outcome with peer pressure if the players are sufficiently sensitive to disadvantageous inequity such that they are willing to reduce it by punishing the other players. Indeed, although it can be costly to punish the other players, subjects can gain some utility by punishing free riders since it is a way to repair a form of injustice.

4. Relaxing assumptions

4.1. Particular case: $c=1$

This section focuses on the case in which $c=1$, i.e. the cost of punishment is equal to the sanction. I assumed in the previous section, that a punisher incurred a cost given by $cY$, $0 < c \leq 1$. I prove that in the case when $c=1$, players are not incited to punish their peers and that defection is a dominant strategy in the first stage of the game.

**Proposition 3**: when the cost of punishment is equal to the sanction (i.e. $c=1$), then cooperation cannot emerge.

**Proof**: If $c=1$, then condition (7) becomes:

$$U_i(Y > 0) - U_i(Y = 0) > 0$$

$$\Leftrightarrow \frac{1}{n-1}[\beta_i(m-1) - \alpha_i(n-m-1)] > 1$$

(12)

This condition is never verified. Players are not incited to punish since punishment cannot reduce disadvantageous inequity relative to the shirkers. The difference of monetary payoffs between each enforcer and the shirkers remains the same. By backward induction, in first stage of the game, the dominant strategy for each player is to defect.

4.2. Magnitude of sanctions

Let us consider now the particular case in which the penalty $Y$ is independent of the number of punishers$^6$. Ostracism can be an example of such sanction since the penalty from being ostracized by one’s peers is independent of the number of excluders. In this case the condition given by equation (7) becomes:

$$U_i(Y > 0) - U_i(Y = 0) > 0$$

$$\Leftrightarrow [\beta_i(m-1) - (n-1) - \alpha_i(n-m)] > 0$$

(13) It is easy to observe that the above condition is never verified. Indeed, if all other players punish, player $i$ is incited to deviate by not punishing since sanctions are independent from the number of punishers. Players anticipate that there will be no punishment in the second stage of the game and all players defect in first stage.

---

$^6$ James S. Coleman (1990) distinguishes incremental sanctioning, in which cost incurred by each sanctioner is small and the effect is also small but the effects are additive and heroic sanctioning, in which the total effect of a sanction occurs through a single actor’s action (for example, to bell the cat in the Aesop’s fable “The mice in council”). The overall difference between these two sanctions lies in the magnitude of the costs incurred by the sanctioners at every stage.
5. Discussion

A comparison of models with and without peer pressure reveals that the availability of punishment supports cooperation. Indeed, if subjects are sufficiently upset by inequity, they are willing to punish the defectors even though this is costly to themselves. One can compare this reaction to the satisfaction for a victim to punish his aggressor (Lindbeck 1997). In the subgame perfect Nash equilibrium of the game, it is a best response to punish defectors in the second stage. By backward induction, the best response in the first stage is to cooperate since threat of punishment is credible.

This analysis also reveals the limits of peer pressure. Indeed, if the cost of punishment is equal to the sanction, then punishment is not a credible threat since it cannot reduce inequity. Moreover, the efficiency of peer pressure also depends on the magnitude of sanctions. If sanctions are independent of the number of punishers such that a defector receives the same sanction whatever be the number of punishers, then peer pressure is ineffective.

While this analysis of peer pressure is illustrative, a number of topics for future research remain. A direction for future research would be to consider the influence of non-monetary sanctions. Non-monetary sanctions are key to the enforcement of implicit agreements and social norms. Kandel and Lazear (1992) explored how non-monetary sanctions (disapproval) operates. Falk, Fehr and Fischbacher (2000) examined the determinants of such informal sanctions by a large number of experiments. Since informal sanctions cannot reduce the payoff differences, inequity aversion cannot explain the efficiency of such peer pressure. This would mean that sanctions would not only be driven by the willingness to reduce payoff differences.

References

CARPENTER J. (1999)”Mutual monitoring in teams : theory and experimental evidence”, working paper Middlebury College


Appendices

Appendix A: non cooperation without sanction

In order to verify that the equilibrium where nobody cooperates is the unique equilibrium under weak conditions, let us consider another proposed equilibrium where all subjects cooperate to produce the good. To verify that the strategy consisting in cooperating supports a Nash equilibrium, I must show that no player is incited to deviate from this proposed equilibrium. Let us consider the utility function of player $i$ when all players cooperate. The payoff of each subject $i$ can be written as:

$U_i(X > 0) : f(n,n) - X$

(A1) If a player $i$ deviates from the equilibrium and does not cooperate, then his payoff is given by:

$U_i(X = 0) : f(n-1,n) - \frac{\beta}{n-1}(n-1)X$

(A2)

When player $i$ deviates from the equilibrium, he saves the cost $X$ for cooperating but experiences now advantageous inequity aversion relative to the $n-1$ other subjects who cooperate. Player $i$ is incited to deviate from the proposed equilibrium if the following condition is verified:

$\beta_i < 1 - \frac{f(n,n) - f(n-1,n)}{X}$

(A3)

Thus if players obey the condition $\beta_i < 1 - \frac{f(n,n) - f(n-1,n)}{X}$, then there is an unique equilibrium with no cooperation.

Appendix B

$$
(1) \frac{dg}{d\alpha_i} = (1-c(n-m)) > 0 \\
(2) \frac{dg}{d\beta_i} = (m-1)c > 0 \\
(3) \frac{dg}{dc} = \beta_i(m-1) - (n-1) - \alpha_i(n-m) < 0 \\
(4) \frac{dg}{dn} = (-\alpha - 1)c < 0 \\
(5) \frac{dg}{dm} = \beta_i c > 0
$$