Abstract

The collective action problem present in team production has been widely discussed in the economics literature, both theoretically and empirically. Based on existing theory, it seems obvious that principals should avoid the simplest form of team production in which the principal claims some share of the value generated and transfers the rest (in equal shares) to the team as compensation. Such a compensation scheme is undermined by the incentive of egoistic workers to shirk. However, in light of recent experimental evidence, it seems unreasonable to assume all workers have purely self-regarding preferences. What if they are inequality averse towards to principal? I use the Fehr-Schmidt model of inequality aversion to show that shirking in this environment is attenuated if workers are inequality averse, and it can be completely eliminated if workers consider their cost of their own efforts when assessing inequality and are sufficiently superiority averse.

JEL Classification: C72, D01, J59,
Keywords: Team production, Social Preferences, Inequality Aversion, Principal, Organization of the firm.
INTRODUCTION

Within the labor economics literature, team production has been one of the most frequently analyzed forms of organizing the firm. The standard result states that, if workers are left alone in the production process, they will always have an incentive to shirk, given that the marginal benefit of each worker depends negatively on the size of the team. In order to solve this so-called “1/n problem”, it has been suggested that some type of monitor or principal is needed (Alchian and Demsetz (1972)). The principal in this case would have to be able to “break the budget”, either in the form of punishment (output being taken away from workers) or reward, so workers would have an incentive to provide the efficient level of effort (Holmstrom (1982)). Moreover, the assumption that workers have purely self-regarding preferences reinforced the notion that extrinsic incentives must be used to create efficiency in teams.

Although self-regarding preferences are a good predictor of behavior in several important social interactions, the relatively recent literature in behavioral and experimental economics has provided robust evidence that people do not care only about their own payoffs in many important situations. These social preferences assume that an agent’s utility is not only a function of the agent’s final result, but may also include the material resources allocated to relevant reference agents (Fehr and Fischbacher (2002)). The most common types of social preferences analyzed in the literature are: 1. Altruism (Andreoni (1989)); 2. Spite (Kirschteiger (1994) and Mui (1995)); 3. Inequality Aversion (Fehr and Schmidt (1999) and Bolton and Ockenfels (2000)); and 4. Reciprocity (Rabin (1993), Falk and Fischbacher (1999), and Charness and Rabin (2002)). As a consequence of these findings, a series of theoretical models have been built trying to incorporate these social preferences into the utility function of agents (either assuming the function to be outcome-based or intention-based) (Dufwenberg and Kirschteiger (1998), Falk and Fischbacher (1999), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), and Charness and Rabin (2002)). Among the outcome-based models of social preferences, Fehr and Schmidt (1999) model of inequality aversion has been the one that has received the most attention in the literature, both from a theoretical and empirical perspective (Lambert et al. (2003), Tricomi et al. (2010), Engelmann (2012)).

This new literature generated a reassessment of the models of team production. So far, the most prominent studies have attempted to derive optimal contracts to solve the collective action problem in this type of labor relation under this new postulation (Bartling and Siemens (2004), Itoh (2004), and Rey-Biel (2007)). This paper will attempt something different. It will show how inequality aversion towards the principal can also implement efficiency in team production under an extremely simple but non-intuitive contract, in which a “rent extracting” principal is added, and all she must do is select in advance the fraction of total output that she wants to receive from workers. Specifically, the model demonstrates how workers choose to provide the efficient level of effort when the principal takes a portion of total output without “breaking the budget”. Effort is higher even though workers in this case receive a smaller portion of total output compared to a situation without a principal.

The first assumption of the model is that agents are inequality averse, following Fehr and Schmidt’s (1999) utility function. In addition to the previously mentioned prominence of this outcome-based function, this utility

---

1 For a discussion about when the traditional models assuming self-regarding preferences predict well and when they do not, see Smith (1998).
function has the advantage of being simple and in most cases capable of rationalizing reciprocal behavior and other similar intention-based behaviors observed in experiments.

Because the main objective is to show how the presence of a principal affects the effort level of a team of inequality averse workers, the model also assumes that the fraction of total output left to workers is divided equally among them, and that they have identical levels of inequality aversion, which is common knowledge to all of them. It is important to notice, however, that in this case workers are only inequality averse towards the principal and not towards each other. The principal is the only reference point of inequality in the production process. This last assumption is justified in two ways. Firstly, theoretical and empirical studies have shown that workers are indeed affected by changes in the incomes of principals (Greenberg (1990), Agell and Lundborg (1995), Bewley (1999), and Dur and Glazer (2006)). Secondly, the experimental evidence on focal points suggests that uniqueness might be a good indicator of what reference point is used by the agents (Schelling (1960)). In a situation where workers are paid equal shares, and the only agent who can get a different share is the principal, we find it plausible to assume that the principal will serve as the reference point of inequality.

As a final assumption, the model considers in one of the cases the possibility that agents take into account their subjective costs of effort when assessing inequality between them and the principal. In an environment where workers do not include this cost in the measurement of inequality, only the monetary payoff of the worker is compared to the payoff of the principal to assess inequality. Although the model also shows the results of this case, it is intuitively plausible to consider the case where the cost of effort is included in the assessment of what is unequal or not.

The paper is organized as follows: section 2 describes the set-up of the model, as well as its respective subsections. It analyzes the behavior of workers when there is no principal, when the principal is added and the inequality averse agents do not take into account their subjective costs of effort in the measurement of inequality, and principal with inequality averse agents who take into account their subjective costs of effort. Section 3 discusses the implications of the results.

2 THE MODEL

Consider a team with either zero or one principal, and \( n \geq 2 \) workers, who are indexed \( i=1,...,n \). If there is a principal, she moves first and chooses to transfer a fraction \( a \in [0, 1] \) of total output to workers in order to maximize her profits; each worker provides a level of effort \( e_i \) to maximize her utility, after observing \( a \). Effort is observable but non-contractible. Let \( c \) be the cost of providing this effort, which is assumed to be quadratic in effort and equal for all workers. These actions determine the monetary outcome \( x \), which is given by the simple production function \( x(e) = \sum_{i=1}^{n} e_i \). The fraction of total output transferred to the agents is divided equally among them, so each

---

2 It is also worth mentioning the turmoil that American Airlines and Delta Airlines’ workers caused in 2003 when they were asked to accept wage cuts while learning that senior executives were earning bonuses.
agent gets $a\sum_{i=1}^{n}e_i/n$. The remainder goes to the principal. The profit function and the utility functions are common knowledge for all players.

2.1. No Principal

In order to serve as the basis for comparison, the model starts analyzing the equilibrium when the team does not have to share total output with a principal. The utility function of each worker $i$ in this case is:

$$U_i = \frac{\sum_{i=1}^{n}e_i}{n} - ce_i^2.$$  

(1)

It is important to emphasize again that workers only suffer inequality aversion towards the principal. In the absence of such a person, there is no inequality aversion and the agents behave in a purely self-regarding manner.

According to this function, workers maximize their utility by providing a level of effort equal to:

$$e_i^* = \frac{1}{2cn}.$$  

(2)

Which falls short of the level of effort that maximizes the collective utility function, given by:

$$e_i^* = \frac{1}{2c}.$$  

(3)

By substituting (2) into the utility function, each agent gets the following in equilibrium:

$$U_i = \frac{(2n-1)}{4cn^2}.$$  

(4)

2.2. “Rent Extracting” Principal with purely self-regarding agents

If the previous result is inefficient, then adding an unproductive principal who just extracts a portion of total production to herself only makes things worse, of course. The principal’s profit function is given by:

$$\pi = (1-a)\sum_{i=1}^{n}e_i.$$  

(5)

Each agent’s utility is then given by:

$$U_i = \frac{a\sum_{i=1}^{n}e_i}{n} - ce_i^2.$$  

(6)

The best-response function of worker “$i$” is:
\[ e^*_i = \frac{a}{2cn}. \] \hspace{1cm} (7)

It is easy to see that (7) is always lower than (2) if \( a < 1 \).

The principal maximizes profit with respect to “\( a \)”, subject to the best-response function of each worker:

\[ \frac{\partial \pi}{\partial a} = \frac{(1-2a)}{2c} = 0, \text{ so the optimal transfer is } a^* = \frac{1}{2}. \] \hspace{1cm} (8)

By substituting (8) into the workers’ best-response functions (7), it is possible to find that each worker provides a level of effort equal to:

\[ e_i = \frac{1}{4cn}. \] \hspace{1cm} (9)

The principals’ profit is equal to:

\[ \pi = \frac{1}{8c}. \] \hspace{1cm} (10)

And the utility of each worker in equilibrium is:

\[ U_i = \frac{(2n-1)}{16cn^2}. \] \hspace{1cm} (11)

These values clearly show that the presence of a “rent extractor” when workers have self-regarding preferences is disastrous for both the production process and the workers, even when compared to the inefficient case analyzed in section 2.1.

2.3. “Rent-Extracting” Principal with inequality averse agents who do not include their subjective cost of effort in their assessment of inequality

Now, agents not only care about their own revenue, but also about the principal’s revenue. They feel disutility by earning either more or less than the principal, with the latter effect being more pronounced. The utility function is given by:

\[ U_i = \frac{a}{n} \sum_{i=1}^{n} e_i - ce_i^2 - \alpha \max \left[ (1-a) \sum_{i=1}^{n} e_i - \frac{a}{n} \sum_{i=1}^{n} e_i ; 0 \right] - \beta \max \left[ \frac{a}{n} \sum_{i=1}^{n} e_i - (1-a) \sum_{i=1}^{n} e_i ; 0 \right]. \] \hspace{1cm} (12)

Where, \( 0 \leq \beta < 1 \), and \( \alpha \geq \beta \), as developed in Fehr and Schmidt (1999). All workers have the same levels of inequality aversion.

Parameter \( \alpha \) captures disadvantageous (or negative) inequality aversion, i.e., the feeling of shame or envy for earning less than the reference agent (the principal in this case). Parameter \( \beta \) captures advantageous (or positive) inequality aversion, which can be represented by the feeling of guilt for earning more than the principal. The function assumes that negative inequality is greater in magnitude than positive inequality. Agents dislike earning
less more than they dislike earning more. The condition $0 \leq \beta < 1$ is also included because if $\beta \geq 1$, then agents would be willing to spend $1 to reduce inequality by exactly $1, which is equivalent to being willing to money away in order to eliminate the inequality. Clearly that is not plausible.

Revenues between the principal and agent “$i$” will be equal when

$$a = \frac{n}{n + 1}. \quad (13)$$

If the revenues are equal, then there is no inequality and the utility function of each agent becomes equal to (6). Condition (13) implies that, as the group size increases, the principal has to transfer a larger fraction of total output in order to generate equal revenues.

Workers feel disadvantageous inequality when:

$$a < \frac{n}{n + 1}. \quad (14)$$

The best-response function of each worker in this case is given by:

$$e^*_i = \frac{a(1 + \alpha + an) - an}{2cn}. \quad (15)$$

Equation (15) is always lower than (7) when (14) holds. Therefore, for any $a$ that creates negative inequality to the workers, they provide a lower level of effort than purely self-regarding agents.

The principal chooses the following transfer:

$$a^* = \frac{2cn + \alpha + 1}{2cn + 2\alpha + 2}. \quad (16)$$

Workers feel advantageous inequality when:

$$a > \frac{n}{n + 1}. \quad (17)$$

The best-response function of each worker is:

$$e^*_i = \frac{a(1 - \beta n - \beta) + \beta n}{2cn}. \quad (18)$$

The principal chooses the following transfer:

$$a^* = \frac{1 - \beta - 2\beta n}{2 - 2\beta - 2\beta n}. \quad (19)$$

**Proposition 1:** When workers suffer from inequality aversion and do not include their subjective cost of effort in their assessment of inequality, the principal never transfers a fraction of output that is sufficient to generate positive inequality.
This is proven by demonstrating that (19) is never simultaneously greater than expression (13) and less than 1. Equation (19) is less than 1 only when \( \beta < \frac{1}{n+1} \), and greater than (13) only when \( \beta > \frac{1}{n+1} \), so both conditions cannot happen at the same time. Consequently, the boundary condition is never satisfied.

Thus, the choice of the principal in this case is always between sharing revenues equally with workers or transferring less to them, which creates negative inequality. The principal selects the latter if (16) gives an optimal residual claim that satisfies (14). The condition for that is:

\[
\alpha < \frac{n-1}{n+1}.
\]  

(20)

In summary, the principal chooses to generate negative inequality (workers earning less) if (20) holds. If it does not hold, she offers (13) and there is no inequality.

**Proposition 2:** When workers suffer from inequality aversion and do not include their subjective cost of effort in their assessment of inequality, as the team increases in size, the likelihood that the principal will share revenues equally decreases.

Expression (20) specifies that the principal will choose to transfer less when the agents’ degree of negative inequality aversion is sufficiently small. But, since \( \frac{\partial \alpha}{\partial n} = \frac{1}{(n+1)^2} > 0 \) from (20), when the group size increases, the interval of negative inequality aversion necessary to make the principal offer an unequal payoff increases as well.

The final levels of effort, output, profits, and utility in each condition are:

(i) If \( \alpha < \frac{n-1}{n+1} \) (negative inequality condition), then:

\[
e_i = \frac{\alpha + 1}{4cn}.
\]  

(21)

\[
\pi = \frac{(\alpha + 1)^2}{8c(\alpha n + \alpha + 1)}.
\]  

(22)

\[
U_i = \frac{(2n-1)(\alpha + 1)^2}{16cn^2}.
\]  

(23)

(ii) If \( \alpha \geq \frac{n-1}{n+1} \) (no inequality condition), then:
\[ e_i = \frac{1}{2c(n+1)}. \]  \hspace{1cm} (24)

\[ \pi = \frac{n}{2c(n+1)^2}. \] \hspace{1cm} (25)

\[ U_i = \frac{(2n-1)}{4c(n+1)^2}. \] \hspace{1cm} (26)

**Proposition 3:** When workers suffer from inequality aversion and do not include their subjective cost of effort in their assessment of inequality, the principal always transfers more to workers than when workers have purely self-regarding preferences.

This is proven by showing that both (13) and (16) are larger than 0.5.

**Corollary 1:** When workers suffer from inequality aversion and do not include their subjective cost of effort in their assessment of inequality, the principal has lower profits than in the case of workers with self-regarding preferences, and these profits decline as the group size increases.

In order to prove the first assertion, it is sufficient to show that expressions (22) and (25) are both smaller than (10). Equation (22) is larger than (10) only if \( \alpha > \frac{n-1}{n+1} \), which is exactly the opposite condition for the principal to choose to generate negative inequality. And for (25) to be larger than (10), the number of workers should be less than 1, which is also an impossibility.

For the second assertion, the conditions for \( \frac{\partial \pi}{\partial n} < 0 \) from equations (22) and (25) are that \( n > 1 \) and \( \alpha > 0 \), which are both true.

**Corollary 2:** When workers suffer from inequality aversion and do not include their subjective cost of effort in their assessment of inequality, they provide a higher level of effort compared to workers with purely self-regarding preferences, but not as high as workers without a principal.

This proposition is corroborated by showing that (21) and (24) are both always greater than (9). Equation (21) is larger than (9) when \( \alpha + 1 > 0 \), which is always true. Similarly, equation (24) is greater than (9) when \( n < 1 \), which is also true.

In summary, compared to the purely self-regarding model, inequality aversion in this case causes the principal to be more generous and workers to exert a higher level of effort. However, this level of effort is still below the efficient one, and workers still have a lower level of utility compared to the no principal case.
2.4. “Rent-Extracting” Principal with inequality averse agents who include their subjective cost of effort in their assessment of inequality

Now, workers subtract their subjective costs of effort from their revenues when they compare their payoffs to the principal’s. The utility function is equal to:

\[
U_i = \frac{a}{n} \sum_{i=1}^{n} e_i - ce_i^2 - \alpha \max \left(1 - a \right) \sum_{i=1}^{n} e_i - \frac{a}{n} \sum_{i=1}^{n} e_i + ce_i^2; 0 \right) - \beta \max \left[ \frac{a}{n} \sum_{i=1}^{n} e_i \right] - ce_i^2 - (1 - a) \sum_{i=1}^{n} e_i, 0 \right].
\] (27)

In this case, payoffs are equal when

\[
e_i = \frac{na + a - n}{c}.
\] (28)

In this case, both \(a\) and effort interact in order to determine whether workers will suffer from negative, positive, or no inequality. In section 2.3, when the principal chooses \(a\), the “type” of inequality does not change, for any given positive level of effort provided by the workers.

The difference between the earnings of the principal and the earnings of worker “\(i\)” in this framework is positively related with effort, which means that an effort higher than (28) causes negative inequality, and an effort lower than (28) causes positive inequality.

The workers’ best-response functions in cases where there is inequality are:

1. Negative Inequality: \(e_i^* = \frac{(a - cn + can + ca)}{2cn(1 + \alpha)}\). \(\) (29)

2. Positive Inequality: \(e_i^* = \frac{(a + \beta n - \beta na - \beta a)}{2cn(1 - \beta)}\). \(\) (30)

If the fraction of total output transferred by the principal is less than \(\frac{n}{n + 1}\), then any positive level of effort creates negative inequality to workers. Within this interval, a level of effort equal to (29) is inferior to providing no effort at all when:

\[a < \frac{can}{can + \alpha + 1}\]. \(\) (31)

Thus, workers provide no effort \((e_i = 0)\) when (31) is satisfied.

If we compare the best-response functions given by expressions (29) and (30) to the ones we found in the previous section (expressions (15) and (18)), we can see that for any given \(a\), the inclusion of the cost of effort in the assessment of inequality reduces the workers’ levels of effort in the presence of negative inequality and increases the levels of effort in the case of positive inequality. It is this increase in effort for positive inequality situations that will

---

3 This expression does require, however, that all workers exert the same level of effort in equilibrium.
drive the main result of this section and generate the main result of the paper, because it may give the principal an incentive to choose a significantly more generous transfers, which will cause workers to provide significantly higher levels of effort out of guilt. And this incentive was not present in the previous section.

The $a$ that maximizes the principal’s profit is the same as (16) and (19) for negative inequality and positive inequality, respectively.

The interval in which workers maximize utility by providing a positive level of effort and one that still creates negative inequality to workers (29) is:

$$\frac{cn}{cn + \alpha + 1} < a < \frac{2n^2 + 2cn^2 - cn}{2n^2 + 2cn^2 + cn + 2n - \alpha - 1}. \quad (32)$$

If $a$ satisfies (32), workers provide a level of effort equal to (29).

The interval in which workers maximize utility by providing a positive level of effort and one that creates positive inequality (30) is:

$$a > \frac{2n^2 - 2\beta n^2 + \beta n}{2n^2 - 2\beta n^2 - \beta n + 2n + \beta - 1}. \quad (33)$$

Similarly, if (33) is satisfied, workers provide a level of effort equal to (30). However, for a transfer greater than (33) the principal’s profit begins to decrease, which means that she will never transfer more than (33).

When $a$ lies within the interval:

$$\frac{2n^2 + 2cn^2 - cn}{2n^2 + 2cn^2 + cn + 2n - \alpha - 1} < a < \frac{2n^2 - 2\beta n^2 + \beta n}{2n^2 - 2\beta n^2 - \beta n + 2n + \beta - 1},$$

workers are not able to maximize utility by providing a level of effort equal to either (29) or (30). Therefore they will choose the level of effort that equalizes the payoffs between them and the principal (expression (28)).

The workers’ best-responses can be summarized as follows: when expression (31) holds they provide no effort; when (32) holds they provide a positive level of effort that creates negative inequality (expression (29)); when (34) holds they provide a positive level of effort that generates no inequality between them and the principal (accounting for the cost of effort) (expression (28)); and when (33) holds, they exert a positive level of effort that creates positive inequality (expression (30)). The next step, now, is to analyze how the principal sets $a$ to maximize her profit.

We have already argued that expression (31) and (33) will never be profit-maximizing for the principal, because the first generates no effort and the second lower profit than no inequality. Therefore, there are only three levels of $a$ that are candidates to maximize the principal’s profit under this environment, and the principal’s decision depends primarily on the degree of inequality aversion that workers have. In the first case, the principal maximizes her profits by choosing to give relatively little, which causes negative inequality to the workers. The fraction $a$ chosen by the principal in this situation will be equal to equation (16) in Case 2.3, namely:

$$a^* = \frac{2cn + \alpha + 1}{2cn + 2\alpha + 2}.$$
For this $\alpha$, workers will exert a level of effort equal to (29). However, this situation will take place only if both $\alpha$ and $\beta$ are small enough. The sufficient condition for the principal to maximize profit at this point is:

$$\frac{(16n^3 - 8n^2)(\alpha n + \alpha + 1)}{(\alpha + 1)} < \frac{(2n^2 - 2\beta n^2 - \beta n + 2n + \beta - 1)^2}{(1 - \beta)}. \quad (35)$$

This condition tells when the principal maximizes profit by setting $a$ equal to (16).

If (35) does not hold, then workers will not suffer inequality, and will provide a level of effort equal to (28). Now, the principal has to choose between two different transfers at this point:

$$a^* = \frac{2n^2 - 2\beta n^2 + \beta n}{2n^2 - 2\beta n^2 - \beta n + 2n + \beta - 1}. \quad (36)$$

Or:

$$a^* = \frac{2n + 1}{2n + 2}. \quad (37)$$

When the workers’s best-response functions are substituted into the principal’s profit function, it is possible to observe that in each interval the principal faces a quadratic profit function. Because equation (36) corresponds to the frontier between “no inequality” and “positive inequality” for the workers, whereas equation (37) is the fraction that maximizes profits when there is no inequality, the principal will choose the lower of these two values. Specifically, the principal will transfer (37) only if:

$$\beta > \frac{2n^2 - 1}{2n^2 + n - 1}. \quad (38)$$

Therefore, the principal transfers the most when workers have a sufficiently high degree of positive inequality aversion. A summary of the three possible outcomes follows:

(i) If $\beta < \frac{2n^2 - 1}{2n^2 + n - 1}$ and $\frac{(16n^3 - 8n^2)(\alpha n + \alpha + 1)}{(\alpha + 1)} \leq \frac{(2n^2 - 2\beta n^2 - \beta n + 2n + \beta - 1)^2}{(1 - \beta)}$, then:

$$a^* = \frac{2n^2 - 2\beta n^2 + \beta n}{2n^2 - 2\beta n^2 - \beta n + 2n + \beta - 1}.$$

$$e_i = \frac{n}{c(2n^2 - 2\beta n^2 - \beta n + 2n + \beta - 1)}.$$

$$\pi = U_i = \frac{n^2(2n - 1)(1 - \beta)}{c(2n^2 - 2\beta n^2 - \beta n + 2n + \beta - 1)^2}. \quad (39)$$

(ii) If $\beta \geq \frac{2n^2 - 1}{2n^2 + n - 1}$ (no inequality), then:
\[ a^* = \frac{2n+1}{2n+2}. \]
\[ e_i = \frac{1}{2c}. \]  
\[ \pi = U_i = \frac{n}{4c(n+1)}. \]  

(iii) If \( \frac{(16n^3 - 8n^2)(\alpha n + \alpha + 1)}{(\alpha + 1)} < \frac{(2n^2 - 2\beta n^2 - \beta n + 2n + \beta - 1)^2}{(1 - \beta)} \) (negative inequality), then:
\[ a^* = \frac{2\alpha n + \alpha + 1}{2\alpha n + 2\alpha + 2}. \]
\[ e_i = \frac{1}{4cn}. \]
\[ \pi = \frac{(\alpha + 1)}{8c(\alpha n + \alpha + 1)}. \]
\[ U_i = \frac{(2n - 1)(\alpha + 1)}{16cn^2}. \]

**Proposition 4:** When workers suffer from inequality aversion and take into account their subjective costs of effort when assessing inequality, if the degrees of inequality aversion are such that the principal creates negative inequality to workers, they will provide the same level of effort as in the case of self-regarding preferences, even though the workers receive a larger fraction of total output.

The proposition is proven by observing that (9) and (43) are equal to each other. Given that (9) is smaller than the level of effort when workers are inequality averse but do not take into account their cost of effort, it is possible to conclude that, when workers have a sufficiently small degree of inequality aversion, workers who take into account the cost of effort when assessing inequality work less than those who do not.

**Corollary 3:** When workers suffer from inequality aversion and take into account their subjective costs of effort when assessing inequality, if the degrees of inequality aversion are such that the principal creates negative inequality to workers, the principal will have lower profits than when workers do not consider this cost of effort.

Given that the fraction of total output transferred by the principal in both cases are the same when \( \alpha \) and \( \beta \) are sufficiently small, but workers provide less effort, naturally profits will be lower. This is verified by showing that (44) is less than (22). The condition for that is when \( \alpha + 1 > 0 \), which is always satisfied.
Thus, it is possible to assert that the principal who faces inequality averse workers with a sufficiently small degree of inequality aversion will always prefer to have the ones who do not take into account the cost of effort when assessing inequality.

**Proposition 5:** When workers suffer from inequality aversion and take into account their subjective costs of effort when assessing inequality, they provide the efficient level of effort if the degree of positive inequality aversion is sufficiently high and the number of workers is greater than 2.

When $\beta \geq \frac{2n^2 - 1}{2n^2 + n - 1}$, workers offer a level of effort equal to the one that maximizes the collective utility function of the workers (3), and as a result this unusual “rent-extracting” contract completely solves the free-rider problem, because effort does not decrease with group size. If the equation is rearranged to show $n$ in terms of $\beta$, the expression also gives the optimal size of the team, given its degree of positive inequality aversion. For example, if $\beta = 0.89$, the collective action problem will be solved up to a group of size 4; and if $\beta$ increases to 0.96, it will support a group of 12.

**Corollary 4:** When workers suffer from inequality aversion and take into account their subjective costs of effort when assessing inequality, their degrees of positive inequality are sufficiently high, and team size is greater than 2, they are better off by having a “rent-extracting” principal than no principal.

When (38) is satisfied, this assertion is corroborated by showing that (42) is greater than (4). The condition for that is $n^2 - 2n - 1 > 0$, which holds for any $n > 2$. However, it is also possible for workers to have a higher level of utility than in the no principal case by providing a level of effort equal to (39). The condition for that is given by the following expression:

$$\beta > \frac{2n^4 + 6n^3 - 2n^2 - 3n + 1 + \left(2n^2 \sqrt{n^4 - 2n^3 - n^2 + n}\right)}{4n^4 + 4n^3 - 3n^2 - 2n + 1}. \quad (46)$$

The principal will always have greater profit in this case compared to the previous inequality aversion case. Nevertheless, the principal’s profit will be larger than the case of workers with self-regarding preferences up to a certain point. This suggests that, given a certain level of positive inequality aversion, the principal will prefer these workers up to a specific team size. If the group continues to increase, the principal will shift her preferences towards having self-regarding workers. This statement is corroborated by analyzing the condition under which equation (40) is greater than equation (10). It happens when:

$$\beta > \frac{4n^4 - 2n^3 + 2n^2 - 3n - 1 + 2\sqrt{8n^6 - 16n^5} + 4n^4 - 2n^2 + 3n}{4n^4 + 4n^3 - 3n^2 - 2n + 1}. \quad (47)$$
Corollary 5: When workers suffer from inequality aversion and take into account their subjective costs of effort when assessing inequality, as the group size increases, the likelihood that workers will be better off by having a principal increases as well.

For a second time, \( \beta \) has to be greater than a certain value for workers to prefer having a principal. Nevertheless, the threshold \( \beta \) given by equation (46) decreases with group size\(^4\), which means that, given a certain level of positive inequality aversion, it is more likely for workers to be included in this interval. As a consequence of this property, it is possible to calculate the exact value of \( \beta \) that will guarantee a preference for a “rent-extracting” principal under this condition, regardless of the size of the group. As stated in Proposition 5 and Corollary 4, the preference for a principal only occurs when the number of workers is greater than 2. By substituting 3 for \( n \) in (46), the value found for \( \beta \) is about 0.54. As this value decreases with group size, as long as workers have a degree of positive inequality aversion greater than 0.54, they will always prefer to have a principal whose only job is to extract rent from them.

Corollary 6: When workers suffer from inequality aversion and take into account their subjective costs of effort when assessing inequality, the principal will prefer workers with a sufficiently high level of positive inequality aversion.

This is proven by showing that (40) and (42) are always greater than the profit functions in all other cases. This result implies that the principal would not prefer inequality averse workers to selfish workers if they do not use the cost of effort in the inequality comparison, but she would prefer inequality averse workers to selfish workers if they take this cost into consideration, and if these workers are sufficiently averse towards positive inequality.

3 DISCUSSION

This paper demonstrated several implications of how the addition of inequality aversion into a model of team production with “rent-extracting” principal can alter the incentives of both parties. Perhaps one of the most important features of the model was to show that the shirking problem in teams is solved not only because workers suffer from inequality aversion, but also due to what is included in the comparison of payoffs. Although purely material inequality increases effort, it is welfare inequality (taking into account the subjective cost of effort) that solves the collective action problem. It happens when workers have a significant aversion towards positive inequality in comparison with the principal. The sequential aspect of the interaction explains the intuition: the principal anticipates this feeling and offers a very generous fraction of total output to workers, who now receive a larger material payoff, but highly dislike having this advantageous inequality. This induces them to increase effort

---

\(^4\) The proof showing how \( \frac{\partial \beta}{\partial n} < 0 \) for equation (46) is excessively long.
considerably in order to equalize welfare. This interaction causes workers to prefer having a principal, even though they face a lower marginal benefit of effort than when no principal is present.

Although this is a simple model, it motivates the analysis of how firms can be organized differently. Specifically, it shows how inequality aversion may create incentives for firms to assign a principal who pretty much just extracts rents form workers. As long as these workers are sufficiently superiority averse and the group size is not too large, the principal will have an incentive to act in a very generous manner in order to induce reciprocity from workers in the form of higher effort levels. Even though a contract like that would be disastrous in the absence of social preferences, it solves to collective action problem here. In a way, the intuition for the results of the model is similar to Becker’s Rotten Kid Theorem. Moreover, social preferences may actually contribute to a larger role for the principal when the production process is similar to team production. It is not the only way of getting efficient results in the presence of social preferences, but it is one possible way.

Naturally, the results of the model have to be corroborated under more relaxing assumptions and different forms of organization, such as heterogeneity of preferences, incomplete information, and when the principal is also productive, for example. The addition of more realistic assumptions of human behavior that go beyond the traditional *homo economicus*, and that are also observed in experimental studies, is a fruitful exercise to understand the implications that they may have to important social interactions involving the firm and the best ways of organizing it.

4 REFERENCES


