

# Rational Sabotage in Cooperative Production with Heterogeneous Agents\*

Carmen Bevia and Luis C. Corchón

## Abstract

We present a model of cooperative production in which rational agents might carry out sabotage activities that decrease output. We provide necessary and sufficient conditions for the existence of a Nash equilibrium without sabotage. It is shown that the absence of sabotage in equilibrium depends on the interplay between technology, relative productivity of agents and the degree of meritocracy. In particular we show that, *ceteris paribus*, meritocratic systems give more incentives to sabotage than egalitarian systems.

**KEYWORDS:** cooperative production, sharing rules, sabotage

---

\*We thank the editor, two referees, Francisco Alcalá, Clara Eugenia García, Shashikanta Nandebam, Inés Macho-Stadler, David Pérez-Castrillo, and Clara Ponsati for their helpful comments. The first author acknowledges financial support from grants SEJ2005-01581/ECON-FEDER, 2005SGR-00454 and the support of the Barcelona Economics program (CREA). The second author acknowledges financial support from BEC2002-02194.

## 1. Introduction

In this paper we consider a model where rational agents have the option to carry out sabotage activities. Traditionally, sabotage has been associated with capitalism -at least since the days of the Luddite revolt- and/or with envy (see Mui (1995)). In this paper we show that sabotage can arise as a rational action under cooperative production, where output is entirely distributed to workers. An example of how sabotage may arise in this framework follows.

Two people are collecting grapes. Andy collects white grapes -whose quantity is denoted by  $R_1$ - and Beth collects red grapes, whose quantity is denoted by  $R_2$ . These grapes are transformed into wine -denoted by  $X$ - according to the production function  $X = (R_1 + R_2)^{1/2}$ . The quantity of wine allocated to each worker - $C_1$  and  $C_2$  respectively- is given by the *Proportional Sharing Rule*, i.e.

$$C_i = \frac{R_i}{R_1 + R_2} (R_1 + R_2)^{1/2}, \quad i = 1, 2.$$

For future reference we notice that this sharing rule is meritocratic, in the sense that it allocates wine depending on relative inputs. Suppose that when an unexpected event forces Beth to leave,  $R_1 = R_2 = 50$ . Thus,  $X = 10$ ,  $C_1 = C_2 = 5$ . Choices for Andy are to remain faithfully devoted to collecting grapes, in which case he would obtain, say, 21 extra units or to destroy the crop assembled by Beth and pretend that somebody stole it.<sup>1</sup> In the first case his consumption of wine is

$$C_1 = \frac{71}{121} (121)^{1/2} \simeq 6,45.$$

In the second case, Andy's consumption of wine is

$$C_1 = \frac{50}{50} (50)^{1/2} \simeq 7,07.$$

Therefore, if Andy is rational and only cares about his wine consumption, he will destroy Beth's crop. Suppose now that the sharing rule is *Egalitarian*, i.e.

$$C_i = \frac{(R_1 + R_2)^{1/2}}{2}, \quad i = 1, 2.$$

---

<sup>1</sup>Andy can explain that his output is just 50 by saying that he spent the remaining time chasing the thief.

We notice that this rule is not meritocratic at all, in the sense that it allocates wine irrespective of relative inputs. In this case, faithful work yields to Andy  $C_1 = 5.5$  and sabotage  $C_1 = 3.5$ , i.e. sabotage is not a rational action.

What is going on in this example? When an agent decides to sabotage there are two effects. On the one hand, output falls reflecting the decrease in the quantity of input supplied by both the saboteur and the agent who has been sabotaged. This is bad from the saboteur's point of view because there is less to be distributed. We will call this the *Production Effect*. Notice that the importance of this effect depends on the returns embodied in the production function. Since returns reflect the degree of congestion among inputs, the magnitude of the production effect depends inversely on the degree of congestion. On the other hand, the relative ranking of the saboteur rises and that is good for him. We will call this the *Distribution Effect*. The importance of this effect depends on how meritocratic the sharing rule is; for instance in the egalitarian sharing rule this effect does not exist. When the rule is meritocratic and there is congestion, the distribution effect may dominate and sabotage is a rational action, as in the case of the proportional sharing rule above.<sup>2</sup>

The model of cooperative production is presented in Section 2. In order to make the model tractable we make a number of simplifications. First we assume that the total quantity of labor supplied by each agent is fixed. Thus, labor can be spent on the production of an intermediate input (by exerting effort) or on the destruction of the inputs of other agents (sabotage). This assumption is made in order to focus attention on the choice between productive and sabotage activities. It is appropriated when length of working time is fixed exogenously by law, custom, etc. Second, we assume that the production function and the sharing rule are symmetric in inputs. However the model is not symmetric because agents have different productivities.<sup>3</sup> This paper studies the existence and properties of Pareto Efficient Nash equilibria,

---

<sup>2</sup>A related example may help to further understand this relationship. Suppose that a group of athletes are running in an event. Runners can devote their energies either to running or to stepping on other people shoes, i.e. to sabotage. One would expect that the occurrence of sabotage depends on how efforts translate into run/sabotage, how runners are rewarded -the more meritocratic the reward, the more sabotage- and the degree of congestion i.e. if runners form a compact pack or they are scattered.

<sup>3</sup>Agents must be heterogeneous because with identical agents all sharing rules yield equal distribution so the question of the distribution of the output, does not arise. Moreover, given that the egalitarian sharing rule yields incentives not to sabotage, this sharing rule is preferable to any other in this case.

which, under our assumptions, is the Nash equilibrium with no sabotage.

In Section 3 we present a necessary condition for sabotage not to arise in a Nash equilibrium that says that for each pair of agents, either the relative capability of sabotage with respect to productive activities is small relative to the relative productivity, or, the degree of meritocracy should be bounded. This bound depends on the degree of congestion, the relative productivity of agents, the relative capability of agents between sabotage and productive activities and the complementarity or substitutability among inputs. This result is recorded as Proposition 1. We also show how this condition can be simplified because for a given agent, say  $i$ , the fulfillment of the necessary condition with respect to another agent, say  $j$ , implies the fulfillment of the necessary condition of  $i$  with respect to any agent less productive than  $j$  (Proposition 2). However, the fulfillment of the necessary condition for some  $i$  does not imply the fulfillment of necessary conditions for any other agent, see Example 1. In our Example 2 we explore the role of the different variables in the necessary condition in the case where the production function is of the CES type and the sharing rule is a convex combination of the proportional and the egalitarian sharing rules. An interesting implication of our analysis is that, for more than two agents with different productivities, an increase in the degree of complementarity among inputs decreases the level of meritocracy compatible with no sabotage. In the limit (perfect complementarity), the unique sharing rule compatible with no sabotage is the egalitarian sharing rule.

In Section 4 we study the existence of a Nash equilibrium with no sabotage. As it happens with the necessary condition, we have two cases. When the capabilities of sabotage with respect to productive activities are small, no sabotage is a Nash equilibrium (Proposition 3). Furthermore, when these relative capabilities are sufficiently small, zero sabotage is the unique Nash equilibrium (Proposition 4). However, when these capabilities are not small we need an extra assumption that is stronger -but with the same flavor- than the necessary condition in order to guarantee that no sabotage is a Nash equilibrium (Proposition 5). Example 3 shows the necessity of this new assumption.<sup>4</sup>

Finally, Section 5 comments our assumptions and suggests further research.

---

<sup>4</sup>Our results contrast with Holmstrom (1982) where efficiency can not be achieved. There are several differences between our model and Holmstrom's, the most important being that in our model agents produce an intermediate input that is contractible. As has been shown by Nandeibam (2002) in a model where sabotage is not possible, efficiency might arise in equilibrium if there are contractible intermediate inputs. Thus, our results and Nandeibam's show the importance of contractible intermediate inputs.

Let us now comment on other papers dealing with sabotage. As far as we know Nalebuff and Stiglitz (1983) were the first to acknowledge that "In the competitive system.... there are... rewards from engaging in destructive activity" (id. p. 40).<sup>5</sup> If saboteurs are identical, we have the following contributions. In the case of a profit maximizing firm, Lazear (1989) showed that, if sabotage is possible, large differences in salaries become dysfunctional. In his model, agents are paid according to the position achieved in a contest.<sup>6</sup> Auriol, Friebel and Pechlivanos, (1999) considered a Principal-Agent model where the former can not commit so she has to offer collectively oriented incentive schemes in order to avoid sabotage. Konrad (2000) considered a model of Rent-Seeking where the effort of an agent reduces rival's performance by sabotaging her activities. He shows that, in equilibrium, sabotage disappears if the number of agents is sufficiently large. Heterogeneous agents have been considered by three different strands of literature: Experimental (Harbring, Irlenbusch, Krakël, and Selten (2004) and Harbring and Irlenbusch (2005)), Management (Chen (2003) and Yumoto (2003)) and Political economy (Skaperdas and Grofman (1995)). All these papers identify conditions under which some of the agents are more likely to engage and/or to receive sabotage.<sup>7</sup> However, none of these papers deals with the main focus of our paper which is to identify conditions on the sharing rule such that sabotage does not arise in equilibrium.

In all the models referenced above agents compete directly so the existence of sabotage is quite natural. In a model of cooperative production where agents at the same time compete (the distribution effect) and cooperate (the production effect) things are more involved. In our model the possibility of sabotage depends on: (1) The degree of meritocracy (as in profit-maximizing firms), (2) the relative productivity of agents (under identical agents the relative productivity of an agent is just one divided by the number of agents, hence the role of the later in rent-seeking models), (3) the relative capability of agents between sabotage and productive activities, (4) the relationship among inputs

---

<sup>5</sup>Itoh (1991) and Macho-Stadler and Pérez-Castrillo (1993) analyze the polar case where cooperation among agents is possible.

<sup>6</sup>When some aspects of performance are rewarded but others are not, other types of dysfunctional behavior are possible, see Holmstrom and Milgrom (1991). This literature is surveyed in Gibbons (1998) and Prendergast (1999).

<sup>7</sup>For instance Skaperdas and Grofman (1995) considered a model in which candidates may focus the campaign on positive or negative aspects, the role of the latter being similar to sabotage. They showed that in equilibrium the leading candidate is less likely to engage in negative campaigning.

(complementarity or substitutability) and (5) the degree of congestion. Thus, our analysis of the necessary and sufficient conditions for absence of sabotage produces a picture where all the causes of sabotage considered before have a role and new causes emerge, i.e. points (4) and (5) above which are associated with the technology of the firm.

## 2. The Model

The setting is one of cooperative production, see, e.g., Sen (1966), Fabella (1988) and Roemer and Silvestre (1993) for examples and applications. There are  $n$  agents. The input provided by agent  $i$  is denoted by  $R_i \in \mathbb{R}_+$ . Let  $X$  be total output. The production function is written  $f(R_1, \dots, R_n) = F(\sum_{k=1}^n \psi(R_k))$ . We will assume that  $f$  is  $\mathcal{C}^1$ , concave, strictly increasing in all its arguments and  $\psi$  is homogeneous of degree  $\rho \in \mathbb{R}$  with  $\rho \leq 1$ . This form generalizes the well known CES form. Total output is shared among agents by means of a sharing rule, i.e. a list of functions  $s_i : \mathbb{R}_+^n \rightarrow [0, 1]$ ,  $i = 1, \dots, n$  such that  $\sum_{i=1}^n s_i(R_1, \dots, R_n) = 1$  for all  $(R_1, \dots, R_n) \in \mathbb{R}_+^n$ . We assume that for all  $i \in \{1, \dots, n\}$ ,  $s_i(R_1, \dots, R_n) = s(R_i, \sum_{k \neq i} R_k)$ , where  $s(\cdot)$  is a  $\mathcal{C}^1$  function, non decreasing on  $R_i$ , non increasing in  $\sum_{k \neq i} R_k$ , homogeneous of degree zero, and such that  $s(R_i, \sum_{k \neq i} R_k) > 0$  if  $R_i > 0$ . The homogeneity assumption ensures that shares do not depend on how inputs are measured. An example of a class of sharing rules fulfilling these conditions is:

$$s_i(R_1, \dots, R_n) = \frac{\alpha R_i}{\sum_{k=1}^n R_k} + \frac{1 - \alpha}{n}, \quad \alpha \in [0, 1], i = 1, \dots, n. \quad (2.1)$$

This family of sharing rules is parametrized by  $\alpha$ . If  $\alpha = 0$  we get the egalitarian sharing rule and if  $\alpha = 1$ , we have the proportional sharing rule. The parameter  $\alpha$  is a measure of how relative effort is valued and thus measures the degree of meritocracy. The interested reader can find in Moulin (1987) and Pfingsten (1991) other examples of sharing rules fulfilling our conditions.

Agents care only about their own consumption. As we remarked in the Introduction, the quantity of labor time is fixed. An agent, say  $i$ , can divide her working time, denoted by  $T$ , between productive labor, denoted by  $l_i^P$  and sabotage activities. Let  $l_{ij}$  be the quantity of labor allocated by  $i$  to sabotage the input of agent  $j$ . The time constraint reads  $T = l_i^P + \sum_{j \neq i} l_{ij}$ . The input provided by agent  $i$  depends on her own productive effort,  $l_i^P$ , the amount of time devoted by the remaining agents to sabotaging the input of

$i$  and a parameter,  $\lambda_i > 0$ , reflecting the productivity of  $i$ . Formally,  $R_i = \lambda_i R(l_i^P, l_{1i}, \dots, l_{(i-1)i}, l_{(i+1)i}, \dots, l_{ni})$  where  $R(\cdot)$  is a  $\mathcal{C}^1$  function such that

$$\frac{\partial R}{\partial l_i^P} > 0 \text{ and } \frac{\partial R}{\partial l_{ji}} < 0.$$

Without loss of generality we will assume that  $\lambda_1 \leq \dots \leq \lambda_n$ . Given these elements, we describe the *sabotage game* as follows: for each agent  $i$ , a strategy is the time devoted to sabotage activities, i.e. the vector  $l_i = (l_{i1}, l_{i2}, l_{i(i-1)}, l_{i(i+1)}, l_{in})$ . Time devoted to productive activities is determined by the constraint  $l_i^P = T - \sum_{j \neq i} l_{ij}$ . By  $l_{-i}$  we denote the vector  $(l_1, \dots, l_{i-1}, l_{i+1}, \dots, l_n)$ . For each agent  $i$ , given a vector of strategies  $(l_i, l_{-i})$ , the payoff function is given by

$$\begin{aligned} \pi_i(l_i, l_{-i}) &\equiv s_i(R_1(l_i, l_{-i}), \dots, R_n(l_i, l_{-i})) f(R_1(l_i, l_{-i}), \dots, R_n(l_i, l_{-i})) \text{ where} \\ R_j(l_i, l_{-i}) &\equiv \lambda_j R(T - \sum_{j \neq i} l_{ji}, l_{1j}, \dots, l_{(j-1)j}, l_{(j+1)j}, \dots, l_{nj}), \quad j = 1, \dots, n. \end{aligned}$$

A *Nash equilibrium of the sabotage game*, denoted by *NE*, is a vector of strategies  $(l_1, \dots, l_n)$  such that for all agent  $i$ ,  $\pi_i(l_i, l_{-i}) \geq \pi_i(l'_i, l_{-i})$  for all  $l'_i$ .

We postpone until Section 4 the problem of the existence of a NE. In Section 3 below we concentrate on the implications of the necessary condition guaranteeing that no agent has incentives to engage in sabotage.

### 3. A Necessary Condition for No Sabotage

If all working time is devoted to productive activities,  $(l_i, l_{-i}) = (0, 0)$ . Define  $R_j^0 \equiv \lambda_j R(0, 0)$  for agent  $j$ , and let  $R^0$  denote the vector of inputs evaluated at the point of zero sabotage, that is,  $R^0 = (R_1^0, \dots, R_n^0)$ . If no agent has incentive to sabotage when all other agents do not sabotage, it must be that  $\forall i, j$ ,  $\frac{\partial \pi_i(0,0)}{\partial l_{ij}} \leq 0$ , where

$$\begin{aligned} \frac{\partial \pi_i(0,0)}{\partial l_{ij}} &= f(R^0) \left( \frac{\partial s_i(R^0)}{\partial R_j} \frac{\partial R_j(0,0)}{\partial l_{ij}} - \frac{\partial s_i(R^0)}{\partial R_i} \frac{\partial R_i(0,0)}{\partial l_i^P} \right) \\ &+ s_i(R^0) \left( \frac{\partial f(R^0)}{\partial R_j} \frac{\partial R_j(0,0)}{\partial l_{ij}} - \frac{\partial f(R^0)}{\partial R_i} \frac{\partial R_i(0,0)}{\partial l_i^P} \right) \end{aligned}$$

If this requirement is not fulfilled, all NE imply sabotage. Let

$$M_{ij} \equiv - \frac{\frac{\partial R_j(0,0)}{\partial l_{ij}}}{\frac{\partial R_i(0,0)}{\partial l_i^P}} = - \frac{\lambda_j \frac{\partial R(0,0)}{\partial l_{ij}}}{\lambda_i \frac{\partial R(0,0)}{\partial l_i^P}}.$$

$M_{ij}$  is a measure of the relative impact of change in inputs induced by an infinitesimal reallocation of  $i$ 's labor from productive to sabotage activities towards agent  $j$  evaluated at the point of zero sabotage. Thus,  $M_{ij}$  is a measure of the power of destruction versus production capabilities. From our assumptions it follows that  $M_{ij} > 0$ . Let

$$M \equiv -\frac{\frac{\partial R(0,0)}{\partial l_{ij}}}{\frac{\partial R(0,0)}{\partial l_i^P}},$$

which is independent of  $i$  and  $j$ . Thus, we can write

$$M_{ij} = \frac{\lambda_j}{\lambda_i} M.$$

By our assumptions on the production function,

$$\frac{\partial f(R^0)}{\partial R_j} = \frac{\partial F(y^0)}{\partial y} \psi'(R_j^0), \text{ where } y \equiv \sum_{k=1}^n \psi(R_k).$$

Since  $\psi$  is homogeneous of degree  $\rho$ ,  $\psi'(R_j^0) = \lambda_j^{\rho-1} \psi'(R(0,0))$ . Thus,

$$\frac{\partial f(R^0)}{\partial R_j} = \left(\frac{\lambda_j}{\lambda_i}\right)^{\rho-1} \frac{\partial f(R^0)}{\partial R_i}.$$

Using the definition of  $M_{ij}$  and dividing by  $f(R^0)$ , the necessary condition reads

$$-\frac{\partial s_i(R^0)}{\partial R_j} M_{ij} - \frac{\partial s_i(R^0)}{\partial R_i} + s_i(R^0) \frac{\partial f(R^0)}{\partial R_i} \frac{1}{f(R^0)} \left(-\left(\frac{\lambda_j}{\lambda_i}\right)^{\rho-1} M_{ij} - 1\right) \leq 0.$$

Or,

$$-\frac{\partial s_i(R^0)}{\partial R_j} \left(\frac{\lambda_j}{\lambda_i}\right) M - \frac{\partial s_i(R^0)}{\partial R_i} + s_i(R^0) \frac{\partial f(R^0)}{\partial R_i} \frac{1}{f(R^0)} \left(-\left(\frac{\lambda_j}{\lambda_i}\right)^\rho M - 1\right) \leq 0.$$

Let  $z \equiv \sum_{k \neq i} R_k$ . Because our assumptions on the sharing rules,

$$\frac{\partial s_i(R^0)}{\partial R_j} = \frac{\partial s(R_i^0, z)}{\partial z},$$

and since the sharing rule is homogenous of degree zero,

$$\frac{\partial s(R_i^0, z)}{\partial R_i} R_i^0 + \frac{\partial s(R_i^0, z)}{\partial z} z = 0.$$



Thus,

$$-\frac{\partial s_i(R^0)}{\partial R_j} = \frac{\partial s_i(R^0)}{\partial R_i} \frac{R_i^0}{\sum_{k \neq i} R_k^0} = \frac{\partial s_i(R^0)}{\partial R_i} \frac{\lambda_i}{\sum_{k \neq i} \lambda_k}.$$

Using this relation, the first order condition can be written as

$$\left( \frac{\lambda_j}{\sum_{k \neq i} \lambda_k} M - 1 \right) \frac{\partial s_i(R^0)}{\partial R_i} \frac{1}{s_i(R^0)} \leq \frac{\partial f(R^0)}{\partial R_i} \frac{1}{f(R^0)} \left( \left( \frac{\lambda_j}{\lambda_i} \right)^\rho M + 1 \right).$$

Multiplying by  $R_i^0$ , the above inequality implies that

$$\left( \frac{\lambda_j}{\sum_{k \neq i} \lambda_k} M - 1 \right) \frac{\partial s_i(R^0)}{\partial R_i} \frac{R_i^0}{s_i(R^0)} \leq \frac{\partial f(R^0)}{\partial R_i} \frac{R_i^0}{f(R^0)} \left( \left( \frac{\lambda_j}{\lambda_i} \right)^\rho M + 1 \right). \quad (3.1)$$

Now we have two cases. If  $M \leq \frac{\sum_{k \neq i} \lambda_k}{\lambda_j}$ , the inequality (3.1) always holds.

However, if  $M > \frac{\sum_{k \neq i} \lambda_k}{\lambda_j}$ ,

$$\frac{\partial s_i(R^0)}{\partial R_i} \frac{R_i^0}{s_i(R^0)} \leq \frac{\partial f(R^0)}{\partial R_i} \frac{R_i^0}{f(R^0)} \left( \left( \frac{\lambda_j}{\lambda_i} \right)^\rho M + 1 \right) \left( \frac{\sum_{k \neq i} \lambda_k}{\lambda_j M - \sum_{k \neq i} \lambda_k} \right). \quad (3.2)$$

Next Proposition summarizes our previous arguments:

**Proposition 1.** *If zero sabotage is a Nash equilibrium of the sabotage game, then either*

1. *The possibilities of destruction are small compared with the relative productivity of all other agents, i.e.,  $M \leq \left( \sum_{k \neq i} \lambda_k \right) / \lambda_j$ , or*
2. *if  $M > \left( \sum_{k \neq i} \lambda_k \right) / \lambda_j$ , in the point of zero sabotage, the elasticity of the share with respect to the input of agent  $i$  is bounded by the elasticity of the production function with respect to the input of agent  $i$ , multiplied by a factor that depends on the possibilities of destruction and the relative productivity of agents (see (3.2) above).*

The elasticity of the share with respect to the input of agent  $i$  can be interpreted as the degree of meritocracy because it measures how one's share responds to one's efforts. The elasticity of the production function with respect to the input of agent  $i$ , can be interpreted as the inverse degree of congestion because it measures how output responds inputs. Thus Proposition 1 says that

when the possibilities of destruction are not small the degree of meritocracy is bounded above by a factor that depends on the degree of congestion, the possibilities of destruction and the relative productivity of agents.

One might think that under reasonable circumstances, it will be possible to simplify the necessary conditions, as it is done in Principal-Agent problems. We may hope that the fulfillment of the necessary condition for some  $i$  implies the fulfillment of necessary conditions for all agents. Or, we may hope that for a given agent, say  $i$ , the fulfillment of the necessary condition with respect to another agent, say  $j$ , implies the fulfillment of the necessary condition of  $i$  with respect to any agent less productive than  $j$ . Next example shows that the first hope is not warranted.

**Example 1.** Let  $n = 2$ ,  $\lambda_1 = 1$ ,  $\lambda_2 = 2$ ,  $s_i(R_1, R_2) = \frac{\alpha R_i}{R_i + R_j} + \frac{1-\alpha}{2}$  and choose  $R(\cdot)$  such that  $M = 2$ . The production function has constant elasticity of substitution,  $X = (\sum (R_i)^\rho)^{\frac{1}{\rho}}$ , with  $\rho \leq 1$  and  $r \leq 1$ . The necessary condition for no sabotage reads:

$$\frac{4\alpha}{3(3-\alpha)} \leq \frac{r}{2^\rho + 1^\rho} (2^{\rho+1} + 1) \text{ for agent 1 and}$$

$$\frac{4\alpha}{3(\alpha+3)} \leq \frac{r2^\rho}{2^\rho + 1^\rho} \left( \frac{1^\rho}{2^\rho} 2 + 1 \right) = \frac{r(2+2^\rho)}{2^\rho + 1} \text{ for agent 2.}$$

For  $\rho = 1$ , (inputs are perfect substitutes) these conditions read:

$$\alpha \leq \frac{15r}{4+5r} \text{ and } \alpha \leq \frac{3r}{1-r}.$$

The first (resp. second) inequality holds trivially when  $r > 0.4$  (resp.  $r > 0.25$ ). We see that  $15r/(4+5r) \leq 3r/(1-r)$  if and only if  $0.1 \leq r$ . Thus for  $r \in (0.1, 0.25)$  the fulfillment of the first inequality implies the fulfillment of the second but for  $r \in (0, 0.1)$  is the other way around.

The interpretation of this example is that if production is not very responsive to inputs, i.e.  $r$  is small, sabotage comes from the most productive agent using her superior capabilities to sabotage the other agent, given the scarce impact of her efforts on output. When  $r$  becomes larger the comparative advantage of the most productive agent is to devote all her time to production activities and he becomes the target of sabotage.

Fortunately, our second conjecture is valid.

**Proposition 2.** Given an agent  $i$ , if the inequality (3.2) holds for  $l$ , it holds for all  $j$  such that  $\lambda_j < \lambda_l$ .

**Proof.** To prove this, it is enough to show that if  $\lambda_j < \lambda_l$  then

$$\left( \left( \frac{\lambda_l}{\lambda_i} \right)^\rho M + 1 \right) \left( \frac{\sum_{k \neq i} \lambda_k}{\lambda_l M - \sum_{k \neq i} \lambda_k} \right) < \left( \left( \frac{\lambda_j}{\lambda_i} \right)^\rho M + 1 \right) \left( \frac{\sum_{k \neq i} \lambda_k}{\lambda_j M - \sum_{k \neq i} \lambda_k} \right), \quad (3.3)$$

which is equivalent to show that

$$(\lambda_l - \lambda_j) > \left( \left( \frac{\lambda_j}{\lambda_i} \right)^\rho - \left( \frac{\lambda_l}{\lambda_i} \right)^\rho \right) \sum_{k \neq i} \lambda_k - \left( \lambda_l \left( \frac{\lambda_j}{\lambda_i} \right)^\rho - \lambda_j \left( \frac{\lambda_l}{\lambda_i} \right)^\rho \right) M. \quad (3.4)$$

Since  $\lambda_j < \lambda_l$ ,  $\left( \frac{\lambda_l}{\lambda_j} \right)^\rho$  is increasing in  $\rho$ , and since  $\rho \leq 1$ , it follows that  $\frac{\lambda_l}{\lambda_j} \geq \left( \frac{\lambda_l}{\lambda_j} \right)^\rho$ , thus,

$$\lambda_l \left( \frac{\lambda_j}{\lambda_i} \right)^\rho - \lambda_j \left( \frac{\lambda_l}{\lambda_i} \right)^\rho \geq 0. \quad (3.5)$$

And since  $M > \left( \sum_{k \neq i} \lambda_k \right) / \lambda_j$ ,

$$\left( \lambda_l \left( \frac{\lambda_j}{\lambda_i} \right)^\rho - \lambda_j \left( \frac{\lambda_l}{\lambda_i} \right)^\rho \right) M > \left( \left( \frac{\lambda_j}{\lambda_i} \right)^\rho - \left( \frac{\lambda_l}{\lambda_i} \right)^\rho \right) \sum_{k \neq i} \lambda_k, \quad (3.6)$$

which implies that the right hand side of inequality (3.4) is negative while the left hand side is positive. Thus, inequality (3.4) holds. ■

To conclude this section, the following example, which is a generalization of Example 1, highlights the role of the different variables in the condition (3.2).

**Example 2.** Let us assume that: (i) The sharing rule has the form

$$\frac{\alpha R_i}{\sum_{k=1}^n R_k} + \frac{1 - \alpha}{n}, \quad \text{with } 0 \leq \alpha \leq 1. \quad (3.7)$$

(i) The production function is of CES type,

$$X = \left( \sum_{i=1}^n (R_i)^\rho \right)^{\frac{r}{\rho}} \quad \text{with } \rho \leq 1 \text{ and } r \leq 1. \quad (3.8)$$

Under i) and ii) we have the following:

$$\frac{\partial s_i(R^0)}{\partial R_i} \frac{R_i^0}{s_i(R^0)} = \frac{\alpha \lambda_i \sum_{k \neq i} \lambda_k}{(\sum_{k=1}^n \lambda_k)^2 \left( \frac{\alpha \lambda_i}{\sum_{k=1}^n \lambda_k} + \frac{1-\alpha}{n} \right)}; \quad (3.9)$$

$$\frac{\partial f(R^0)}{\partial R_i} \frac{R_i^0}{f(R^0)} = \frac{r \lambda_i^\rho}{\sum_{k=1}^n \lambda_k^\rho}. \quad (3.10)$$

By Proposition 2 we can restrict our attention to the necessary condition for agent  $i$  not to sabotage agent  $n$ . Thus, inequality (3.2) can be written as

$$\frac{\frac{\alpha \lambda_i}{\sum_{k=1}^n \lambda_k}}{\frac{\alpha \lambda_i}{\sum_{k=1}^n \lambda_k} + \frac{1-\alpha}{n}} \leq \frac{r(\lambda_n^\rho M + \lambda_i^\rho)}{\sum_{k=1}^n \lambda_k^\rho} \left( \frac{\sum_{k=1}^n \lambda_k}{\lambda_n M - \sum_{k \neq i} \lambda_k} \right). \quad (3.11)$$

We now study the effect of the variables in the two sides of (3.11): We see that the effects of  $M$ , and  $r$  are what we expect from intuition, i.e. an increase in  $M$  -the relative power of destruction with respect to production- makes harder the fulfillment of (3.11) and an increase of  $r$  -the responsiveness of output with respect to inputs- makes easier the fulfillment of (3.11). But the effects of  $\rho$  and  $\lambda$ 's are complex, even though some interpretation of what is going on is possible.

*Effect of  $\rho$ .* To analyze the effect of  $\rho$  assume first that  $n > 2$ , and there are at least an agent  $k$  such that  $\lambda_k < \lambda_{n-1}$ . Let us restrict our attention to inequality (3.11) for  $i = n - 1$  and  $j = n$ . Notice that  $\rho$  only affects the expression

$$\frac{r(\lambda_n^\rho M + \lambda_{n-1}^\rho)}{\sum_{k=1}^n \lambda_k^\rho}. \quad (3.12)$$

This expression reflects the effect of sabotage from  $n - 1$  to  $n$  on output. It reflects the opportunity cost of sabotage. It can be shown that (3.12) is increasing in  $\rho$ , and it goes to zero when  $\rho$  goes to minus infinity (perfect complementarity) (see Lemma 2 in the Appendix for details). Thus, in the limit, the unique level of meritocracy compatible with no sabotage is  $\alpha = 0$  (even if we have constant returns to scale,  $r = 1$ ). The interpretation of this result is that when inputs approach perfect complementarity, output depends only on the input provided by the least productive agent, so there are incentives to sabotage any other agent because output will be unaffected.

When  $n = 2$  the situation is different, because a reduction in the input provided by agent 1 -coming either from a sabotage of 2 into 1 or from the distraction

of resources implied by sabotage of agent 1- reduces production. As a consequence, sharing rules can be more meritocratic when there are two agents only. Formally, the expression (3.12) goes to  $r$  when  $\rho \rightarrow -\infty$ . So in the limit inequality (3.11) can be written as:

$$\alpha \leq \frac{r(\lambda_1 + \lambda_2)^2}{2\lambda_1\lambda_2(M - 1) + r(\lambda_2^2 - \lambda_1^2)}. \quad (3.13)$$

It can be shown that the right hand side depends on the relative productivity of agents, and is U-shaped with a unique minimum at  $\lambda_1/\lambda_2 = (M - r - 1) / (M + r - 1)$ . This form reflects the trade-off of sabotage for agent 1: When agent 2 is very productive, sabotage has no noticeable impact on the relative ranking so it is a waste of time. And when both agents are almost identical, sabotage affects output which is bad from the saboteur interests. Thus meritocracy takes the highest possible value in one of the extremes of  $\lambda_1/\lambda_2$ .

*Effect of  $\lambda$ 's.* It is difficult to say something in general about the effect of productivities on (3.11). Let us simplify this inequality by assuming perfect substitution, i.e.  $\rho = 1$ . Then, (3.11) reads

$$\frac{\alpha \frac{\lambda_i}{\sum_{k=1}^n \lambda_k}}{\frac{\alpha \lambda_i}{\sum_{k=1}^n \lambda_k} + \frac{1-\alpha}{n}} \leq \frac{r(\lambda_j M + \lambda_i)}{\lambda_j M - \sum_{k \neq i} \lambda_k}. \quad (3.14)$$

- *Effect of  $\lambda_i$  :* From the previous equation, we see that when  $\lambda_i \rightarrow \infty$  the LHS of (3.14) tends to one but the RHS tends to infinite, thus the inequality holds. In this case, the large productivity of  $i$  yields incentives to concentrate her activities in production matters. Also when  $\lambda_i \rightarrow 0$  the LHS of (3.14) tends to zero but the RHS tends to a positive number and thus (3.14) also holds. This is more intriguing but it can be explained by noting that when  $\lambda_i$  is close to 0, the distribution effect, i.e. the change in  $i$ ' share measured here by the LHS of (3.14), is negligible so sabotage does not pays off. Finally, it is clear that the necessary condition may not hold for intermediate values of  $\lambda_i$  where the interplay between the production and the distribution effect can produce complicated patterns.

- *Effect of  $\lambda_j$  :* When  $\lambda_j \rightarrow \infty$  the LHS of (3.14) tends to zero but the RHS tends to  $r$ , thus (3.14) holds. Again this can be explained by saying that in this case the large productivity of  $j$  makes the distribution effect to be nil. When  $\lambda_j \rightarrow \left( \sum_{k \neq i, j} \lambda_k \right) / (M - 1)$  (which is the lower bound in the case we are analyzing) the LHS of (3.14) tends to a finite number but the RHS tends

to infinite and thus (3.14) also holds. In this case what happens is that we get close to the case where  $M \leq \left(\sum_{k \neq i} \lambda_k\right) / \lambda_j$  in which the necessary condition always holds. Again, the necessary condition may not hold for intermediate values of  $\lambda_j$  for reasons identical to those explained before.

- Effect of  $\sum_{k \neq i, j} \lambda_k$  : When  $\sum_{k \neq i, j} \lambda_k \rightarrow 0$  (3.14) reads

$$\frac{\alpha \frac{\lambda_i}{\lambda_i + \lambda_j}}{\alpha \frac{\lambda_i}{\lambda_i + \lambda_j} + \frac{1-\alpha}{n}} \leq \frac{r(\lambda_j M + \lambda_i)}{\lambda_j(M-1)},$$

which may or may not hold. In this case we are back to the case of two agents where the necessary condition may or may not hold, depending on the interplay between  $\lambda_i$  and  $\lambda_j$ . Finally, when  $\sum_{k \neq i, j} \lambda_k \rightarrow (M-1)\lambda_j$  the LHS of the above inequality tends to a positive number but the RHS tends to infinite and thus the inequality is satisfied. The explanation is identical to the case where  $\lambda_j \rightarrow \left(\sum_{k \neq i, j} \lambda_k\right) / (M-1)$ . Again, the necessary condition may not hold for intermediate values of  $\sum_{k \neq i, j} \lambda_k$  for reasons identical to those explained before.

*Effect of  $\alpha$ .* An increase in  $\alpha$  makes the LHS of (3.11) larger. For given values of all other variables, this makes harder the fulfillment of this condition.

The necessary condition yields an upper bound on the degree of meritocracy compatible with no sabotage. This is very clear when all agents are identical, because the left hand side of (3.11) simplifies to  $\alpha$  and the condition gives us the bound directly. Let us work out the non symmetrical case. Define

$$q_{ij} \equiv \frac{r(\lambda_j^p M + \lambda_i^p)}{\sum_{k=1}^n \lambda_k^p} \left( \frac{\sum_{k=1}^n \lambda_k}{\lambda_j M - \sum_{k \neq i} \lambda_k} \right). \quad (3.15)$$

Now, condition (3.11) can be written as

$$\alpha \left( \frac{\lambda_i}{\sum_{k=1}^n \lambda_k} + q_{ij} \left( \frac{1}{n} - \frac{\lambda_i}{\sum_{k=1}^n \lambda_k} \right) \right) \leq \frac{q_{ij}}{n}. \quad (3.16)$$

Notice that for some agents,  $\frac{\lambda_i}{\sum_{k=1}^n \lambda_k} < q_{ij} \left( \frac{\lambda_i}{\sum_{k=1}^n \lambda_k} - \frac{1}{n} \right)$  and so (3.16) holds for any  $\alpha$ . However, at least, for all agents whose productivity is less than the average -i.e.  $\frac{1}{n} \geq \frac{\lambda_i}{\sum_{k=1}^n \lambda_k}$ - the left hand side of (3.16) is positive and thus it can be written as

$$\alpha \leq \frac{\frac{q_{ij}}{n}}{\frac{\lambda_i}{\sum_{k=1}^n \lambda_k} + q_{ij} \left( \frac{1}{n} - \frac{\lambda_i}{\sum_{k=1}^n \lambda_k} \right)}. \quad (3.17)$$

Let  $S$  be the set of agents for whom  $\frac{\lambda_i}{\sum_{k=1}^n \lambda_k} > q_{ij} \left( \frac{\lambda_i}{\sum_{k=1}^n \lambda_k} - \frac{1}{n} \right)$ . For any  $i \in S$  define

$$\alpha_{ij} \equiv \frac{\frac{q_{ij}}{n}}{\frac{\lambda_i}{\sum_{k=1}^n \lambda_k} + q_{ij} \left( \frac{1}{n} - \frac{\lambda_i}{\sum_{k=1}^n \lambda_k} \right)}, \quad (3.18)$$

and let  $\alpha_i \equiv \min_{j \neq i} \alpha_{ij}$ . Then, the maximum degree of meritocracy compatible with no sabotage is  $\min(1, \min_{i \in S} \alpha_i)$ . By Proposition 2 for each agent  $i$  the most restrictive bound is the one towards the most productive agent, that is,  $\alpha_i = \alpha_{in}$  for all  $i = 1, \dots, n-1$ , and  $\alpha_n = \alpha_{n(n-1)}$ . Furthermore, notice that for each agent  $i$ ,  $\alpha_{ij} < 1$  if and only if  $q_{ij} < 1$ . Otherwise, any degree of meritocracy is compatible with no sabotage. In our discussion on the effect of  $\rho$  we have proved that  $q_{(n-1)n}$  is increasing in  $\rho$ , by (3.18) as  $\rho$  decreases the maximum degree of meritocracy compatible with no sabotage decreases. The most favorable case for meritocracy is the case of perfect substitutes inputs ( $\rho = 1$ ). In this case,

$$q_{ij} \equiv \frac{r(\lambda_j M + \lambda_i)}{\lambda_j M - \sum_{k \neq i} \lambda_k},$$

which is less than one if and only if

$$r < \frac{\lambda_j M - \sum_{k \neq i} \lambda_k}{\lambda_j M + \lambda_i}.$$

In the Cobb-Douglas case, it is easily calculated that

$$\lim_{\rho \rightarrow 0} q_{ij} = \frac{1}{n} r (M + 1) \left( \frac{\sum_{k=1}^n \lambda_k}{\lambda_j M - \sum_{k \neq i} \lambda_k} \right),$$

which is less than one if and only if

$$r < \frac{n(\lambda_j M - \sum_{k \neq i} \lambda_k)}{(M + 1) \sum_{k=1}^n \lambda_k}.$$

## 4. A Sufficient Condition for No Sabotage.

In this section we study under what conditions a Nash equilibrium with no sabotage exists. To make the problem tractable, we assume that the individual's input is given by

$$\begin{aligned} R(l_i, l_{-i}) &= \max(T - \sum_{j \neq i} l_{ij} - K \sum_{j \neq i} l_{ji}, 0), \\ R_i(l_i, l_{-i}) &= \lambda_i R(l_i, l_{-i}), \quad i \in \{1, \dots, n\} \end{aligned} \quad (4.1)$$

Where the parameter  $K$  is a positive constant. Notice that in this case, if agent  $i$  uses one unit of labor time to sabotage the input of agent  $j$ , he reduces his input in one unit and the input of agent  $j$  in  $K$  units. Thus,

$$-\frac{\frac{\partial R(l_i, l_{-i})}{\partial l_{ij}}}{\frac{\partial R(l_i, l_{-i})}{\partial l_i^P}} = K \text{ for all } (l_i, l_{-i}) \text{ such that } T - \sum_{j \neq i} l_{ij} - K \sum_{j \neq i} l_{ji} \geq 0.$$

Next lemma simplifies the search for an equilibrium without sabotage because it characterizes the best response of  $i$  when the rest of agents is not sabotaging anyone.

**Lemma 1.** *Let  $l_i = (l_{i1}, \dots, l_{ii-1}, l_{ii+1}, \dots, l_{in})$  be a best response for agent  $i$  to  $l_{-i} = 0$ , then if  $\lambda_j \leq \lambda_k$ , and  $l_{ij} > 0$ ,  $R_k(l_i, l_{-i}) \leq R_j(l_i, l_{-i})$  and consequently  $l_{ik} \leq l_{ij}$ .*

**Proof.** Notice first that if  $l_i$  is a best response to  $l_{-i} = 0$ , then  $T - Kl_{ij} \geq 0$  for all  $j$ . Suppose not, that is, suppose that there is an agent  $j$  such that  $T - Kl_{ij} < 0$ . Then agent  $i$  can decrease the time dedicated to sabotaging agent  $j$  up to a point such that  $T - Kl'_{ij} = 0$ . Thus agent  $i$  will increase her input without affecting the input of the other agents, which implies that she will be better off and will contradict that  $l_i$  is a best response against  $l_{-i} = 0$ . Suppose that  $R_k(l_i, l_{-i}) > R_j(l_i, l_{-i})$ , let  $\hat{l}_i$  be such that  $\hat{l}_{il} = l_{il}$  for all  $l \neq j, k$ ,  $\hat{l}_{ij} = l_{ij} - \varepsilon_j$ ,  $\hat{l}_{ik} = l_{ik} + \varepsilon_k$  where  $\varepsilon_k > 0$ ,  $\varepsilon_j > 0$ ,  $\lambda_k \varepsilon_k = \lambda_j \varepsilon_j$  and  $\hat{l}_{ij} \geq 0$ . Thus,  $R_k(l_i, l_{-i}) + R_j(l_i, l_{-i}) = R_k(\hat{l}_i, l_{-i}) + R_j(\hat{l}_i, l_{-i})$  and  $R_i(\hat{l}_i, l_{-i}) > R_i(l_i, l_{-i})$ , which implies that  $s_i(\hat{l}_i, l_{-i}) \geq s_i(l_i, l_{-i})$ , so the share of this agent does not decrease. Finally, let us see that the production increases. Recall that  $f(R_1, \dots, R_n) = F(\sum_{j=1}^n \psi(R_j))$  with  $\psi$  homogeneous of degree  $\rho \leq 1$ . We distinguish two cases. First, suppose that  $\rho > 0$ . In this case  $\psi$  is increasing and concave, thus  $\psi(R_i(\hat{l}_i, l_{-i})) > \psi(R_i(l_i, l_{-i}))$ ,  $\psi(R_k(l_i, l_{-i})) - \psi(R_k(\hat{l}_i, l_{-i})) \leq$



$\psi(R_j(\hat{l}_i, l_{-i}) - \psi(R_j(l_i, l_{-i}))$ , and since  $f$  is strictly increasing in all its arguments,  $F$  is strictly increasing in  $\sum_{j=1}^n \psi(R_j)$ , therefore  $f(R_1(\hat{l}_i, l_{-i}), \dots, R_n(\hat{l}_i, l_{-i})) > f(R_1(l_i, l_{-i}), \dots, R_n(l_i, l_{-i}))$ . Secondly, suppose that  $\rho < 0$ . In this case  $\psi$  is decreasing and convex, thus  $\psi(R_i(\hat{l}_i, l_{-i})) < \psi(R_i(l_i, l_{-i}))$ ,  $\psi(R_k(l_i, l_{-i})) - \psi(R_k(\hat{l}_i, l_{-i})) \geq \psi(R_j(\hat{l}_i, l_{-i}) - \psi(R_j(l_i, l_{-i}))$ , and since  $f$  is strictly increasing in all its arguments,  $F$  is strictly decreasing in  $\sum_{j=1}^n \psi(R_j)$ , therefore  $f(R_1(\hat{l}_i, l_{-i}), \dots, R_n(\hat{l}_i, l_{-i})) > f(R_1(l_i, l_{-i}), \dots, R_n(l_i, l_{-i}))$ . In both cases the production increases which implies that  $\pi_i(\hat{l}_i, l_{-i}) > \pi_i(l_i, l_{-i})$ . ■

Lemma 1 implies that if an agent is not sabotaging the most productive agent, he is not sabotaging anyone.

In the last section we proved that if  $M \leq \left(\sum_{k \neq i} \lambda_k\right) / \lambda_j$  for all  $i, j$ , the necessary condition for no sabotage holds. Notice that, given our assumption on the individual's input,  $M = K$ . The next Proposition shows that if  $K \leq \left(\sum_{k \neq i} \lambda_k\right) / \lambda_j$  for all  $i, j$ , zero sabotage is a Nash equilibrium. The intuition is that since the damage that agents can inflict on each other is small, sabotage does not pay off.

**Proposition 3.** Assume  $K \leq \left(\sum_{k \neq i} \lambda_k\right) / \lambda_j$  for all  $i, j$ . Then  $l_{ij} = 0$  for all  $i, j \in \{1, \dots, n\}$  is a Nash equilibrium.

**Proof.** Let us see that for each agent  $i$  the best response to  $l_{-i} = 0$  is  $l_i = 0$ . Suppose on the contrary that the best response to  $l_{-i} = 0$  involves positive sabotage by agent  $i$ . Let  $l_i = (l_{i1}, \dots, l_{ii-1}, l_{ii+1}, \dots, l_{in})$  be such that  $l_{ij} > 0$  for some  $j$ . By Lemma 1,  $l_{in} > 0$ . Define  $R_j^l \equiv \lambda_j R(l_i, 0)$  for all agent  $j$ , and let  $R^l$  the vector of inputs evaluated at the point  $(l_i, 0)$ , that is,  $R^l = (R_1^l, \dots, R_n^l)$ . As we have shown in the proof of Lemma 1,  $T - Kl_{ij} \geq 0$  for all  $j$ . Let us see first that there is at least one agent  $j$  such that  $T - Kl_{ij} > 0$ . Suppose not, then  $T - \sum_{j \neq i} l_{ij} = T - T(n - 1)/K$ . Since  $K \leq \left(\sum_{k \neq i} \lambda_k\right) / \lambda_j$  for all  $i, j$ ,  $K \leq n - 1$ , which implies that  $T - T(n - 1)/K \leq 0$ . Thus  $R_i^l = 0$ . But this can not be the best response of agent  $i$  to  $l_{-i} = 0$ . Since nothing is produced, the payoff of this agent is zero. By reducing the sabotage activities he will get a positive payoff. Therefore  $\sum_{j \neq i} R_j^l > 0$ . Finally, let us see that for all  $l_i$  such that  $\sum_{j \neq i} R_j^l > 0$ ,  $\frac{\partial \pi_i(l_i, 0)}{\partial l_{in}} < 0$ , and thus  $l_{in} = 0$ , which by Lemma 1 implies that  $l_{ij} = 0$  for all  $j$ . By the definition of  $\pi_i$ ,

$$\frac{\partial \pi_i(l_i, 0)}{\partial l_{in}} = \frac{\partial s_i(R^l)}{\partial l_{in}} f(R^l) + s_i(R^l) \frac{\partial f(R^l)}{\partial l_{in}}.$$

From (4.1) we have that  $\frac{\partial R_n(l_i, 0)}{\partial l_{in}} = -K\lambda_n$ , and  $\frac{\partial R_i(l_i, 0)}{\partial l_{in}} = -\lambda_i$ . Since the production function is strictly increasing in all its arguments, it follows that

$$\frac{\partial f(R^l)}{\partial l_{in}} = -\frac{\partial f(R^l)}{\partial R_n} K\lambda_n - \frac{\partial f(R^l)}{\partial R_i} \lambda_i < 0.$$

Thus, in order to see that  $\frac{\partial \pi_i(l_i, 0)}{\partial l_{in}} < 0$ , it is enough to show that

$$\frac{\partial s_i(R^l)}{\partial l_{in}} = -\frac{\partial s_i(R^l)}{\partial R_n} K\lambda_n - \frac{\partial s_i(R^l)}{\partial R_i} \lambda_i \leq 0. \quad (4.2)$$

If the sharing rule is constant, the above inequality always holds. It also holds if  $l_i$  is such that  $R_i^l = 0$ , because since sharing rules are homogeneous of degree zero,

$$\frac{\partial s_i(R^l)}{\partial R_i} R_i^l + \frac{\partial s_i(R^l)}{\partial R_n} \sum_{j \neq i} R_j^l = 0,$$

and since  $\sum_{j \neq i} R_j^l > 0$ ,  $\frac{\partial s_i(R^l)}{\partial R_n} = 0$ . For all other  $l_i$  such that  $R_i^l > 0$ , proving that inequality (4.2) holds, is equivalent to prove that

$$-\frac{\frac{\partial s_i(R^l)}{\partial R_i}}{\frac{\partial s_i(R^l)}{\partial R_n}} \geq \frac{K\lambda_n}{\lambda_i}.$$

Since the sharing rule is homogeneous of degree zero,

$$-\frac{\frac{\partial s_i(R^l)}{\partial R_i}}{\frac{\partial s_i(R^l)}{\partial R_n}} = \frac{\sum_{j \neq i} R_j(l_i, 0)}{R_i(l_i, 0)} = \frac{T \sum_{j \neq i} \lambda_j - K \sum_{j \neq i} \lambda_j l_{ij}}{\lambda_i (T - \sum_{j \neq i} l_{ij})}.$$

Notice that, since  $K \leq \left( \sum_{k \neq i} \lambda_k \right) / \lambda_j$  for all  $i, j$ ,  $\sum_{j \neq i} \lambda_j \geq K\lambda_n$ , and by the order of the agents,  $\lambda_j \leq \lambda_n$ . Thus,

$$\frac{T \sum_{j \neq i} \lambda_j - K \sum_{j \neq i} \lambda_j l_{ij}}{\lambda_i (T - \sum_{j \neq i} l_{ij})} \geq \frac{\lambda_n K (T - \sum_{j \neq i} l_{ij})}{\lambda_i (T - \sum_{j \neq i} l_{ij})} = \frac{\lambda_n K}{\lambda_i},$$

as we wanted to prove. ■

The previous equilibrium is not always unique. If  $\left( \sum_{k \neq i} \lambda_k \right) / \lambda_j \geq K \geq \lambda_i / \lambda_j$  for all  $i, j$  (in particular  $K \geq 1$ ), there are, at least, two kinds of Nash equilibria additional to that with no sabotage:

I) In the first kind, no positive output is produced. This equilibrium can be

sustained with the following strategies: For each  $i$ ,  $l_{ii+1} = T$  ( modulo  $n$ ) and  $l_{ij} = 0$  otherwise. Clearly  $R_i(l_i, l_{-i}) = 0$  for all  $i$ . It is also clear that no agent can deviate profitably.

II) In the second kind (only possible if  $n > 2$ ), only one agent produces a positive input. This equilibrium can be sustained with the following strategies: For each  $i \neq n$ , let  $l_{ii+1} = T$  (modulo  $n - 1$ ) and  $l_{ij} = 0$  otherwise,  $l_{ni} = 0$  for all  $i$ . Clearly  $R_i(l_i, l_{-i}) = 0$  for all  $i \neq n$ , and  $R_n(l_i, l_{-i}) > 0$ . It is also clear that no agent can deviate profitably.

However when  $K < \lambda_i/\lambda_j$  for all  $i, j$ , we can guarantee uniqueness of equilibrium.

**Proposition 4.** *Assume  $K < \lambda_i/\lambda_j$  for all  $i, j$ . Then  $l_{ij} = 0$  for all  $i, j \in \{1, \dots, n\}$  is the unique Nash equilibrium.*

**Proof.** The proof is left to the Appendix.

Let us now consider the case when the possibilities of destruction are large, i.e.  $K > \left(\sum_{k \neq i} \lambda_k\right) / \lambda_j$ . In contrast with the previous case, the necessary condition is no longer sufficient.

**Example 3.** *If in Example 2 we set  $n = 4$ ,  $T = 10$ ,  $M = 5$ ,  $\alpha = 0.5$ ,  $r = 1/3$ ,  $\rho = 1$ ,  $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4$ , the necessary condition holds, but in this case no sabotage is not an equilibrium: the payoff of agent  $i$  in the point of zero sabotage is  $\pi_i(0, 0) = 0.854$  but, for  $l_i = (2, 2, 2)$ ,  $\pi_i(l_i, 0) = 0.992$ . So an increase in sabotage activities pays off.*

The problem in this case is that the share of an agent can increase with sabotage, contrary to what happened when  $K \leq \left(\sum_{k \neq i} \lambda_k\right) / \lambda_j$ , (see 4.2). To guarantee that zero is a Nash equilibrium this effect can not compensate the fall in output caused by sabotage. The following assumption just formalizes this.

A1. For each agent  $i$ , let  $l = (l_i, l_{-i})$  with  $l_{-i} = 0$ ,  $0 \leq l_{ij} \leq \frac{T}{K}$ , for all  $j \neq i$ . Let  $R_i^l = \lambda_i \left(T - \sum_{k \neq i} l_{ik}\right)$ ,  $R_k^l = \lambda_k (T - Kl_{ik})$ , and  $R^l = (R_1^l, \dots, R_n^l)$ . Then,

$$-\frac{\partial s_i(R^l)}{\partial R_n} \frac{R_i^l}{s_i(R^l)} < \frac{\partial f(R^l)}{\partial R_i} \frac{R_i^l}{f(R^l)} \left( \left(\frac{\lambda_n}{\lambda_i}\right)^\rho K + 1 \right) \frac{\lambda_i}{K\lambda_n}. \quad (4.3)$$

A1 is similar to the necessary condition but stronger on two counts. On the one hand it is evaluated not only in the point of zero sabotage but in all

points described above. On the other hand it is a little bit more restrictive since in the point of zero sabotage and by the homogeneity of the sharing rule,

$$-\frac{\partial s_i(R^0)}{\partial R_n} = \frac{\partial s_i(R^0)}{\partial R_i} \frac{\lambda_i}{\sum_{k \neq i} \lambda_k}.$$

Thus, this sufficient condition in the point of zero sabotage reads:

$$\frac{\partial s_i(R^0)}{\partial R_i} \frac{R_i^0}{s_i(R^0)} < \frac{\partial f(R^0)}{\partial R_i} \frac{R_i^0}{f(R^0)} \left( \left( \frac{\lambda_n}{\lambda_i} \right)^\rho K + 1 \right) \frac{\sum_{k \neq i} \lambda_k}{K \lambda_n},$$

and since

$$\frac{\sum_{k \neq i} \lambda_k}{K \lambda_n} \leq \frac{\sum_{k \neq i} \lambda_k}{K \lambda_n - \sum_{k \neq i} \lambda_k},$$

this condition is stronger than the necessary condition (3.2).

We are now prepared to state and prove our next result.

**Proposition 5.** *Assume A1, and  $K > \left( \sum_{k \neq i} \lambda_k \right) / \lambda_j$ , for all  $j$ . Then  $l_{ij} = 0$  for all  $i, j \in \{1, \dots, n\}$  is a Nash equilibrium.*

**Proof.** Let us see that for each agent  $i$  the best response to  $l_{-i} = 0$  is  $l_i = 0$ . Suppose on the contrary that the best response to  $l_{-i} = 0$  involves positive sabotage by agent  $i$ . Let  $l_i = (l_{i1}, \dots, l_{ii-1}, l_{ii+1}, \dots, l_{in})$  be such that  $l_{ij} > 0$  for some  $j$ . By Lemma 1,  $l_{in} > 0$ ,  $l_{in} \geq l_{ij}$  and  $T - Kl_{ij} \geq 0$  for all  $j$ . Define  $R_j^l \equiv \lambda_j R(l_i, 0)$  for all agent  $j$ , and let  $R^l$  the vector of inputs evaluated at the point  $(l_i, 0)$ , that is,  $R^l = (R_1^l, \dots, R_n^l)$ . Let us see that  $\frac{\partial \pi_i(l_i, 0)}{\partial l_{in}} < 0$ , for all  $(l_i, 0)$ , then  $l_{in} = 0$ , which, by Lemma 1 implies that  $l_{ij} = 0$  for all  $j$ . By the definition of  $\pi_i$

$$\frac{\partial \pi_i(l_i, 0)}{\partial l_{in}} = \lambda_i \left( f(R^l) \left( -\frac{\partial s_i(R^l)}{\partial R_n} K \frac{\lambda_n}{\lambda_i} - \frac{\partial s_i(R^l)}{\partial R_i} \right) + s_i(R^l) \left( -\frac{\partial f(R^l)}{\partial R_n} K \frac{\lambda_n}{\lambda_i} - \frac{\partial f(R^l)}{\partial R_i} \right) \right). \quad (4.4)$$

The assumed form of the production function implies that

$$\frac{\partial f(R^l)}{\partial R_n} = \frac{\partial F(y^l)}{\partial y} \psi'(R_n^l), \text{ where } y = \sum_{k=1}^n \psi(R_k). \quad (4.5)$$

Since  $\psi$  is homogeneous of degree  $\rho \leq 1$ ,  $\psi'$  is homogeneous of degree  $\rho - 1$ , hence  $\psi'(R_n^l) = \psi'(\lambda_n(T - Kl_{in})) = \lambda_n^{\rho-1} \psi'(T - Kl_{in})$ . Notice that since  $K >$

$(\sum_{k \neq i} \lambda_k) / \lambda_j$  for all  $i, j$ ,  $K > n - 1$ , Thus,  $T - \sum_{k \neq i} l_{ik} \geq T - (n - 1)l_{in} > T - Kl_{in}$ . Given that  $f$  is strictly increasing in all its arguments, if  $\rho \geq 0$ ,  $\psi'$  is decreasing and  $\frac{\partial F(y^l)}{\partial y} \geq 0$ , if  $\rho < 0$ ,  $\psi'$  is increasing and  $\frac{\partial F(y^l)}{\partial y} \leq 0$ . In any of these cases,

$$\frac{\partial F(y^l)}{\partial y} \lambda_n^{\rho-1} \psi'(T - Kl_{in}) \geq \frac{\partial F(y^l)}{\partial y} \lambda_n^{\rho-1} \psi'(T - \sum_{k \neq i} l_{ik}), \quad (4.6)$$

thus multiplying and dividing the right hand side of 4.6 we have that,

$$\frac{\partial f(R^l)}{\partial R_n} \geq \left( \frac{\lambda_n}{\lambda_i} \right)^{\rho-1} \frac{\partial f(R^l)}{\partial R_i}, \quad (4.7)$$

which implies that

$$-\frac{\partial f(R^l)}{\partial R_n} K \frac{\lambda_n}{\lambda_i} - \frac{\partial f(R^l)}{\partial R_i} \leq \frac{\partial f(R^l)}{\partial R_i} \left( -K \left( \frac{\lambda_n}{\lambda_i} \right)^\rho - 1 \right). \quad (4.8)$$

Since  $\frac{\partial s_i(R^l)}{\partial R_i} \geq 0$ ,

$$-\frac{\partial s_i(R^l)}{\partial R_n} K \frac{\lambda_n}{\lambda_i} - \frac{\partial s_i(R^l)}{\partial R_i} \leq -\frac{\partial s_i(R^l)}{\partial R_n} K \frac{\lambda_n}{\lambda_i}. \quad (4.9)$$

By A1, (4.8), (4.9), and (4.4),  $\frac{\partial \pi_i(l_i, 0)}{\partial l_{in}} < 0$ , as we wanted to prove. ■

## 5. Conclusions

In this paper we have presented a model of cooperative production where agents can sabotage other agent's inputs. We have derived necessary and sufficient conditions to avoid sabotage in a Nash equilibrium. These conditions point out to factors like the technology of the firm, the relative productivity of agents, the relative capability of agents between sabotage and productive activities and the degree of meritocracy. These findings can be used on several counts: To help our understanding of how and why an organization populated by rational agents might be self-destructive (see Genicot and Skaperdas (2002) for a similar point in a model of conflict). To explain why cooperatives tend to offer egalitarian incentive schemes (Kremer (1997) and Priks (2005) and the references there). Or to the design of sharing rules that do not encourage sabotage but are meritocratic.

Our results have been obtained under a number of simplifying assumptions. Thus, it is a fair question to ask what would happen if some of these assumptions are removed. Let us comment on this focussing our attention on two issues: the form of the sharing rules and the assumption that the quantity of leisure is not a choice variable.<sup>8</sup>

**1. More General Sharing Rules:** We may consider sharing rules of the form  $s_i(R_1, \dots, R_n) = s(R_i, \sum_{k=1}^n g(R_k))$  where  $g$  is a non decreasing and concave function. This would allow for sharing rules like  $s_i(R_1, \dots, R_n) = \frac{(R_i)^\lambda}{\sum_{k=1}^n (R_k)^\lambda}$ ,  $\lambda \in [0, 1]$ . Unfortunately, the sufficient conditions considered in this paper for preventing the existence of an equilibrium with sabotage do not work in this case, even if all agents are identical. For instance, when  $K < \lambda_i/\lambda_j$ , there might be equilibria with and without sabotage (see Beviá and Corchón (2003), Example 1) and, in some cases, even if  $K < (\sum_{k \neq i} \lambda_k)/\lambda_j$ , only equilibria with sabotage exists (see Beviá and Corchón (2003), Example 2).

**2. Variable Working Time:** In this paper we assume that the length of the working time is exogenously given. How different is the case where agents can decide the length of working time from the case analyzed in this paper? Clearly, the necessary condition for no sabotage still holds, because any working time has to be distributed optimally between productive and sabotage activities. But in this new framework no sabotage is not necessarily a socially optimal choice: Sharing rules that do not encourage sabotage may also not encourage a high level of effort (e. g. the egalitarian rule). Thus, in some cases, social optimality may require choosing a sharing rule which encourages both effort and sabotage. In consequence, a model with a variable working time requires a considerable departure from the methods developed in this paper that are based in that no sabotage is always socially optimal. This extension should be subject of further research.

---

<sup>8</sup>Other possible extensions are repeated interaction or the consideration of actions like stealing, defending against sabotage or giving information that could damage other people's inputs.

## 6. Appendix.

**Lemma 2.** Let  $n > 2$ ,  $\lambda_k \leq \lambda_{n-2} < \lambda_{n-1} \leq \lambda_n$  with  $k = 1, \dots, n - 3$ , and let  $h(\rho) = (r(\lambda_n^\rho M + \lambda_{n-1}^\rho)) / \sum_{k=1}^n \lambda_k^\rho$ . The function  $h$  is increasing in  $\rho$  and goes to zero when  $\rho$  goes to minus infinity.

**Proof.** First of all notice that the function  $h$  can be written as a function of the relative productivities of all agents with respect to the productivity of agent  $n - 1$ ,

$$h(\rho) = \frac{r\left(\left(\frac{\lambda_n}{\lambda_{n-1}}\right)^\rho M + 1\right)}{\sum_{k=1}^n \left(\frac{\lambda_k}{\lambda_{n-1}}\right)^\rho}.$$

Since  $\lambda_{n-2} < \lambda_{n-1}$ ,  $(\lambda_{n-2}/\lambda_{n-1})^\rho \rightarrow \infty$  as  $\rho \rightarrow -\infty$ . Thus  $h(\rho) \rightarrow 0$  as  $\rho \rightarrow -\infty$ . Finally, let us see that  $h'(\rho) > 0$ .

$$h'(\rho) = r \frac{\sum_{k=1}^n \left(\frac{\lambda_k}{\lambda_{n-1}}\right)^\rho M \left(\frac{\lambda_n}{\lambda_{n-1}}\right)^\rho \ln\left(\frac{\lambda_n}{\lambda_{n-1}}\right) - \left(\left(\frac{\lambda_n}{\lambda_{n-1}}\right)^\rho M + 1\right) \sum_{k=1}^n \left(\frac{\lambda_k}{\lambda_{n-1}}\right)^\rho \ln\left(\frac{\lambda_k}{\lambda_{n-1}}\right)}{\left(\sum_{k=1}^n \left(\frac{\lambda_k}{\lambda_{n-1}}\right)^\rho\right)^2}.$$

The numerator can be written as:

$$\begin{aligned} & \left(\frac{\lambda_n}{\lambda_{n-1}}\right)^\rho \ln\left(\frac{\lambda_n}{\lambda_{n-1}}\right) \left(\sum_{k=1}^n \left(\frac{\lambda_k}{\lambda_{n-1}}\right)^\rho M - \left(\frac{\lambda_n}{\lambda_{n-1}}\right)^\rho M - 1\right) - \\ & \left(\left(\frac{\lambda_n}{\lambda_{n-1}}\right)^\rho M + 1\right) \sum_{k=1}^{n-1} \left(\frac{\lambda_k}{\lambda_{n-1}}\right)^\rho \ln\left(\frac{\lambda_k}{\lambda_{n-1}}\right). \end{aligned}$$

First notice that

$$\left(\left(\frac{\lambda_n}{\lambda_{n-1}}\right)^\rho M + 1\right) \sum_{k=1}^{n-1} \left(\frac{\lambda_k}{\lambda_{n-1}}\right)^\rho \ln\left(\frac{\lambda_k}{\lambda_{n-1}}\right) < 0$$

because for all  $k = 1, \dots, n - 1$ ,  $\ln\left(\frac{\lambda_k}{\lambda_{n-1}}\right) \leq 0$  and since  $\lambda_{n-2} < \lambda_{n-1}$ ,  $\ln\left(\frac{\lambda_{n-2}}{\lambda_{n-1}}\right) < 0$ . Secondly,

$$\sum_{k=1}^n \left(\frac{\lambda_k}{\lambda_{n-1}}\right)^\rho M - \left(\frac{\lambda_n}{\lambda_{n-1}}\right)^\rho M - 1 = \sum_{k=1}^{n-2} \left(\frac{\lambda_k}{\lambda_{n-1}}\right)^\rho M + M - 1 > 0$$

because  $M > \frac{\sum_{k \neq n-1} \lambda_k}{\lambda_n} > 1$ . Therefore  $h'(\rho) > 0$ . ■

**Proof of Proposition 2.**

Suppose we have an equilibrium with positive sabotage,  $(l_1, \dots, l_n)$ , and let  $R^S = (R_1(l_i, l_{-i}), \dots, R_n(l_i, l_{-i}))$ . Then,

**Step 1.** There is at least one agent  $i$  such that  $\sum_{j=1, j \neq i}^n l_{ji} \leq T$ .

Suppose that for all agent  $i$ ,  $\sum_{j=1, j \neq i}^n l_{ji} > T$ . Then, if we sum for all agents,  $\sum_i \sum_{j=1, j \neq i}^n l_{ji} > nT$ , but this is impossible since, because of the time constraint, for all  $j$ ,  $\sum_{i=1, i \neq j}^n l_{ji} \leq T$ .

**Step 2.** There is at least one agent  $i$  such that  $R_i(l_i, l_{-i}) > 0$ .

Suppose that, for all agent  $j$ ,  $R_j(l_i, l_{-i}) = 0$ . Then  $\pi_j(l_i, l_{-i}) = 0$  for all  $j$ . By Step 1 we know that there is an agent  $i$  such that  $\sum_{j=1, j \neq i}^n l_{ji} \leq T$ . Since  $K < \lambda_i/\lambda_j$ , and  $R_i(l_i, l_{-i}) = 0$ , the amount of time devoted to sabotage activities by this agent  $i$  is strictly positive. But this can not be an equilibrium. If this agent reduces her sabotage activities, the total output will be positive and her input positive. Consequently, she will get a positive amount. Therefore, she will be better off.

**Step 3.** There are at least two agents,  $i$  and  $j$ , such that  $R_i(l_i, l_{-i}) \neq R_j(l_i, l_{-i})$ .

Suppose on the contrary that for all  $i$ , and  $j$ ,  $R_i(l_i, l_{-i}) = R_j(l_i, l_{-i})$ . By the assumptions on the sharing rule,  $\pi_i(l_i, l_{-i}) = \frac{1}{n}f(R_1(l_i, l_{-i}), \dots, R_n(l_i, l_{-i}))$ . Suppose agent  $i$  reduces her sabotage activity toward agent  $j$  zero, let  $\hat{l}_i$  denote the new strategy for agent  $i$ . In this case, the input of agent  $i$  will increase by  $\lambda_i l_{ij}$  and the input of agent  $j$  will increase by  $K\lambda_j l_{ij}$ . Since  $K < \lambda_i/\lambda_j$ , and the rest of agents are not affected,  $R_i(\hat{l}_i, l_{-i}) > R_j(\hat{l}_i, l_{-i}) > R_k(\hat{l}_i, l_{-i})$  for all  $k \notin \{i, j\}$ . Thus,  $\pi_i(\hat{l}_i, l_{-i}) \geq \frac{1}{n}f(R_1(\hat{l}_i, l_{-i}), \dots, R_n(\hat{l}_i, l_{-i})) > \frac{1}{n}f(R_1(l_i, l_{-i}), \dots, R_n(l_i, l_{-i}))$ , which implies that agent  $i$  is better off.

**Step 4.** If  $R_j(l_i, l_{-i}) \leq R_i(l_i, l_{-i}) \leq R_k(l_i, l_{-i})$ , and  $T - K \sum_{j \neq i} l_{ji} > 0$ , then  $l_{ik} = 0$  and  $l_{ij} = 0$ .

Suppose first that  $l_{ik} > 0$ . Since total output is decreasing in  $l_{ik}$ , let us see that this can not be an equilibrium because the share of this agent is non increasing in  $l_{ik}$ . That is,

$$-\frac{\partial s_i(R^S)}{\partial R_k} K \lambda_k - \frac{\partial s_i(R^S)}{\partial R_i} \lambda_i \leq 0.$$

If the sharing rule is constant (the egalitarian sharing rule), the above inequality always holds. If it is not, to prove the above inequality is equivalent to prove that

$$-\frac{\frac{\partial s_i(R^S)}{\partial R_i}}{\frac{\partial s_i(R^S)}{\partial R_k}} \geq \frac{K \lambda_k}{\lambda_i}.$$



Since the sharing rule is homogeneous of degree zero and  $K < \lambda_i/\lambda_k$ ,

$$-\frac{\frac{\partial s_i(R^S)}{\partial R_i}}{\frac{\partial s_i(R^S)}{\partial R_j}} = \frac{\sum_{j \neq i} R_j(l_i, l_{-i})}{R_i(l_i, l_{-i})} \geq \frac{R_k(l_i, l_{-i})}{R_i(l_i, l_{-i})} \geq 1 > \frac{K\lambda_k}{\lambda_i}.$$

Thus,  $l_{ik} = 0$ .

Suppose secondly that  $l_{ij} > 0$ . Notice first that if  $T - \sum_{k \neq j} l_{jk} - K \sum_{k \neq j} l_{kj} < 0$ ,  $l_{ij} > 0$  can not be an equilibrium, because agent  $i$  can decrease the time dedicated to sabotaging agent  $j$  without affecting the input of agent  $j$  but increasing her input and total output, which implies that she will be better off. Thus  $T - \sum_{j \neq k} l_{kj} - K \sum_{j \neq k} l_{jk} \geq 0$ . Since total output is decreasing in  $l_{ij}$ , let us see that this can not be an equilibrium because the share of this agent is non increasing in  $l_{ij}$ . That is,

$$-\frac{\partial s_i(R^S)}{\partial R_j} K\lambda_j - \frac{\partial s_i(R^S)}{\partial R_i} \lambda_i \leq 0.$$

If the sharing rule is constant (the egalitarian sharing rule), the above inequality always holds. If it is not, to prove the above inequality is equivalent to prove that

$$-\frac{\frac{\partial s_i(R^S)}{\partial R_i}}{\frac{\partial s_i(R^S)}{\partial R_j}} \geq \frac{K\lambda_j}{\lambda_i}.$$

Since the sharing rule is homogeneous of degree zero and  $K < \lambda_i/\lambda_j$ ,

$$-\frac{\frac{\partial s_i(R^S)}{\partial R_i}}{\frac{\partial s_i(R^S)}{\partial R_j}} = \frac{\sum_{j \neq i} R_j(l_i, l_{-i})}{R_i(l_i, l_{-i})} \geq \frac{R_k(l_i, l_{-i})}{R_i(l_i, l_{-i})} \geq 1 > \frac{K\lambda_j}{\lambda_i}.$$

Thus,  $l_{ij} = 0$ .

**Step 5.** For all agent  $i$ ,  $T - K \sum_{j \neq i} l_{ji} > 0$ .

Suppose on the contrary that there are  $k$  agents such that  $T - K \sum_{j \neq i} l_{ji} \leq 0$ . Suppose that these agents are the first  $k$  agents. By Step 4, these agents do not suffer sabotage from agents  $k + 1$  to  $n$ . Thus  $T - K \sum_{j=1, j \neq i}^k l_{ji} \leq 0$  for all  $i = 1, \dots, k$ . Adding these inequalities for  $i = 1, \dots, k$ , we get:  $kT - K \sum_{i=1}^k \sum_{j=1, j \neq i}^k l_{ji} \leq 0$ . But, by the time constraint,  $\sum_{i=1, i \neq j}^k l_{ji} \leq T$ , and since  $K < 1$ ,  $kT - K \sum_{i=1}^k \sum_{j=1, j \neq i}^k l_{ji} > 0$ .

**Step 6.** There is not an equilibrium with positive sabotage. By the previous steps, we know that for all agent  $i$  except the one with the biggest input,

and for all  $j$ ,  $l_{ij} = 0$ . Thus, if there is an equilibrium with positive sabotage, only the agent with the biggest input is using part of her time in sabotage activities. But, by Proposition 3 the best response to zero sabotage by others is zero sabotage. ■

## References

- [1] Auriol, E., Friebel, G. and Pechlivanos, L. (1999). "Teamwork Management in an Era of Diminishing Commitment." CEPR, #2281.
- [2] Beviá, C., and Corchón, L. (2003). "Rational Sabotage in Cooperative Production." Working Paper UAB-IAE 541.02.
- [3] Chen, K. P. (2003). "Sabotage in Promotion Tournaments." *Journal of Law, Economics, and Organization* 19 (1), 119-139.
- [4] Fabella, R.V. (1988). "Natural Team Sharing and Team Productivity." *Economics Letters* 27, 105-110.
- [5] Genicot, G. and Skaperdas, S. (2002). "Investing in Conflict Management." *Journal of Conflict Resolution*, 46 (1), 154-170.
- [6] Gibbons, R. (1998). "Incentives in Organizations." *Journal of Economic Perspectives*, 12 (4), 115-132.
- [7] Harbring, C., Irlenbusch, M., Krakël, M. and Selten R. (2004). "Sabotage in Asymmetric Contests." Mimeo, University of Cologne and London School of Economics.
- [8] Harbring, C., and Irlenbusch, B. (2005). "How Many Winners Are Good to Have? On Tournaments with Sabotage." IZA Discussion Paper No. 1777, September.
- [9] Holmstrom, B. (1982). "Moral hazard and Teams." *Bell Journal of Economics*, 13, 324-340.
- [10] Holmstrom, B., and Milgrom, P. (1991). "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics and Organization*, 7, Special Issue, 24-52.
- [11] Itoh, H. (1991). "Incentives to Help in Multi-Agent Situations." *Econometrica*, 59 (3), 611-636.
- [12] Konrad, K. A. (2000). "Sabotage in Rent-Seeking Contests." *Journal of Law, Economics and Organizations*, 16 (1), 155-165.
- [13] Kremer, M. (1997). "Why are worker co-operatives so rare?" NBER Working Paper: 6118.0.

- [14] Lazear, E. P. (1989). "Pay Equality and Industrial Politics." *Journal of Political Economy*, 97 (3), 561-580.
- [15] Macho-Staedler, I., and Pérez-Castrillo, D. (1993). "Moral Hazard with Several Agents." *International Journal of Industrial Organization*, 11, 73-100.
- [16] Moulin, H. (1987). "Equal or proportional Division of Surplus and Other Methods." *International Journal of Game Theory*, 16, 161-186.
- [17] Mui, Vai-Lam (1995). "The Economics of Envy." *Journal of Economic Behavior & Organization*, 26, 3, 311-336.
- [18] Nalebuff, B. J., and Stiglitz, J.E. (1983). "Prizes and Incentives: Towards a General Theory of Compensation and Competition." *Bell Journal of Economics*, 14, 21-43.
- [19] Nandeibam, S. (2002). "Implementation in Teams." Working paper, University of Birmingham.
- [20] Prendergast, C. (1999). "The Provision of Incentives in Firms." *Journal of Economic Literature*, 37,7-63.
- [21] Pfingsten, A. (1991). "Surplus Sharing Methods." *Mathematical Social Sciences*, 21, 287-301.
- [22] Priks, M. (2005). "Egalitarian Wage Schedules in Small Firms and in Co-operatives: Rent Seeking at Work." Working Paper Center for Economic Studies, University of Munich.
- [23] Roemer, J., and Silvestre, J. (1993). "The Proportional Solution for Economies with Private and Public Ownership." *Journal of Economic Theory*, 59, 426-444.
- [24] Sen, A. (1966). "Labour Allocation in a Cooperative Enterprise." *Review of Economic Studies*, 33, 361-371.
- [25] Skaperdas, S. and Grofman, B. (1995). "Modeling Negative Campaigning." *American Political Science Review* 89, 49-61.
- [26] Yumoto, Y. (2003). "Who Is a Target of Sabotage?: The Dark Side of Promotion Tournaments." Working Paper, North Carolina State University, July.