CDE

Centre for Development Economics

WORKING PAPER SERIES

Moral Hazard in a Principal-Agent(s) Team

Shasikanta Nandeibam
Centre for Development Economics
Delhi School of Economics

Working Paper No: 17

Centre for Development Economics
Delhi School of Economics
Delhi 110 007 INDIA

Moral Hazard in a Principal-Agent(s) Team*

Shasikanta Nandeibam†
Centre for Development Economics
Delhi School of Economics
Delhi-110007, India

July 1994

*This is an extensively revised version of an earlier paper with the same title.

†I would like to thank Charles Blackorby, David Donaldson, Mukesh Eswaran, Kenneth Hendricks, Ashok Kotwal, William Schworm, Guofu Tan, and John Weymark for helpful comments on earlier versions of this paper. I am also grateful for the hospitality of the Centre for Development Economics, Delhi School of Economics, where this paper was revised while I was a visitor. Of course, I am solely responsible for any errors and omissions.

Abstract

We look at the principal's problem in a principal-agent(s) (possibly more than one agent) moral hazard problem, which, unlike most existing work, does not preclude the principal from active participation in the production process. Also, there is no uncertainty, but joint production, which renders the action of each individual in the production process unobservable, causes the moral hazard problem. The principal and all the agents play a multi-stage extensive game, called the Second Best Game, which determines the set of individuals who actually participate in production along with the output sharing rule they follow[1]. Although the principal is not precluded from active participation in the production process, we characterize a condition that determines whether she actually takes part in production or not. We also draw the following conclusions: (i) the principal need not look for any output sharing rule more sophisticated than those that belong to the class of commonly observed linear or piecewise linear variety; (ii) the principal can completely mitigate moral hazard whenever she does not participate in production; (iii) however, even when the principal does not participate in production at an optimal outcome, she may still be unable to sustain efficiency; (iv) the principal can sustain efficiency if and only if her best option in the First Best situation does not require her participation in production; and (v) there are no significant changes to the results when limited liability is imposed. We argue that most of the results are driven by the deterministic production process and not by the quasilinear form of the utility functions that we use. Hence, as long as the production process is deterministic, most of the results will hold qualitatively even when individuals do not have quasilinear utilities but their utilities remain additively separable and concave.

[1]It will become clear from the text that the actual participants play a noncooperative game in the production process conditional on the output sharing rule they follow.

1 Introduction

One of the most prominent features of existing work on moral hazard in the principal-agent framework is the separation of ownership from labour. In most previous work (*e.g.* Grossman and Hart[7], Harris and Raviv[8], Hart and Holmstrom[9], Holmstrom[10], Rees[15], Ross[16], Shavell[18]) only the agent takes part in production but her action is unobservable and unverifiable because of uncertainty in the production process, and the principal is there only as a passive residual claimant (because she does not take part in production). So, as the residual claimant, the problem of the principal is to design, before production begins, a payment schedule for the agent that depends only on the observable final output. This payment schedule is designed to induce the agent to choose a level of action which will maximize the expected benefit of the principal (from the residual) subject to the incentive and individual rationality constraints.

However, we often observe organizations in which there is a residual claimant who hires the services of other individuals in the production process, designs the output sharing rule, and also takes part in the production process unlike the principal in most principal-agent models. In these organizations, as in most principal-agent relationships, the residual claimant role of the individual who has the right to design the output sharing rule usually stems from her ownership of tangible essential inputs like technology, equipment, capital or even the license to operate the business. Also, the participation of the residual claimant in the production process is often because of her comparative advantage in certain input that is specific to the production technology. Quite a sizable proportion of the so called "self owned and operated" businesses fall under this category, because the self employed owner in such businesses often tend to hire the services of other individuals as well. Such organizations are also quite prevalent in the "small-scale and cottage industries" sector of most less developed countries, where a single individual (or household), because of her ability to make the necessary investments or her ability to acquire the credits for the necessary investments or even her ability to influence the bureaucracy (as is often required) to secure the proper business license, may often start a business that employs herself and others. Also, even though most existing work use the landlord-tenant relationship in the agrarian economies of the less developed countries as a good example of a principal-agent relationship, as Eswaran and Kotwal[6] pointed out, often the landlord not only designs the rule for sharing the crop with the tenant but also makes farm management decisions that are not observable to the tenant.

So there is enough evidence to suggest the coexistence of the two kinds of principal-agent(s) relationship, namely, those in which the principal is only a residual claimant, and those in which the principal is not only a residual claimant but also an active participant in the production process. Then the following question comes to mind immediately. Contrary to the common assumption that the principal's nonparticipation (or participation) is exogenously determined, is it possible that the principal is not precluded from participating in production ex ante in a sizable number of cases, but whether she does participate or not is a choice which she makes? For instance, as Eswaran and Kotwal[6] pointed out, this question often has an affirmative answer in the case of landlord-tenant relationships, because, rather than being an absentee landlord, the landlord can often choose to participate in farming, for example, by making farm management decisions. Once the principal's participation is not precluded, the condition which determines her participation decision will depend on things like her opportunity cost, the degree of complementarity between her action and the actions of the agents, the opportunity costs of the agents, the relative efficiency of any agent whose action can substitute the principal's action, etc. Thus, using a fairly general model with some standard assumptions, we want to characterize the condition which determines the principal's participation decision.

Apart from the principal's participation decision, an equally important question is, how do the optimal output sharing rules look like? We have the following answer to this question. If the principal chooses not to participate, there is an optimal output sharing rule in which each participating agent (we allow the possibility of two or more agents) has a linear payment function with a slope of one. On the otherhand, if the principal's best option is to participate, there is an optimal output sharing rule in which each participating agent has a linear payment function with a slope which lies strictly between zero and one, and moreover, the residual function of the principal is also linear with a slope which lies strictly between zero and one. Although we have quasilinear utility functions, our result on linearity of optimal output sharing rules depends not on this but on another important feature of our model which has to do with the cause of moral hazard. As we want to concentrate on the moral hazard caused by joint production, in contrast to the single-agent case (which precludes the principal's participation) where moral hazard is caused by the presence of uncertainty in production, there are no uncertainties in our production process and moral hazard is a pure joint production phenomenon. Thus, unlike the case with uncertainty where incentive constraints impose conditions on the behaviour of the output sharing rule throughout the support of the distribution of output, because of the absence of uncertainty, incentive constraints impose only a local condition on the behaviour of the output sharing rule around the equilibrium output level and leaves sufficient degrees of freedom to choose the behaviour of the output sharing rule else where. So, in contrast to the case where there is uncertainty in production, qualitatively, our result on linearity of optimal output sharing rules will hold even if the utility functions are no longer quasilinear provided they are still additively separable and concave and there are no uncertainties in production.

The amount of freedom provided by the absence of uncertainty on the behaviour of optimal output sharing rules away from a local neighbourhood of the equilibrium output level also has another important implication for the case where individuals have limited liabilities. Even when individuals have limited liabilities, there is still sufficient freedom to modify the linear optimal output sharing rules in such a way that we get piecewise linear optimal output sharing rules that satisfy the limited liability constraints. This will not be always possible if there are uncertainties in the production process.

We also show that, if the principal does not participate in production, then she cannot do any better even if the actions of the participating agents were made observable. So moral hazard is completely mitigated whenever the principal acts only as a residual claimant. This is because of the fact that, if the principal does not participate in production, then, as in Holmstrom[11], her role is just like the role of the outsider who administers "budget-breaking" incentive schemes. It must be pointed out that, like the result on the linearity of optimal output sharing rules, this result on complete mitigation of moral hazard in the case of nonparticipation by the principal does not depend on the quasilinearity of the utility functions. As long as the utility functions are additively separable and concave and the production process is deterministic, moral hazard can be always completely mitigated in the case of nonparticipation by the principal. This is in sharp contrast to the case where there is uncertainty in production and risk sharing.

On the otherhand, if the principal does participate in production, then she cannot completely mitigate the moral hazard problem, and hence, she can do better if the actions were observable. This is because of the fact that there is an inherent conflict between the principal's role as the residual claimant and her incentive to shirk in the production process.

Thus, the answer to another important question, "Can the principal sustain efficiency?", depends crucially on whether she has to participate in the production process in the full information case (which is the hypothetical situation where all actions are observable) to get the maximum utility. In particular, if the principal does not have to participate in production in the full information case to get the maximum utility, then and only then can she sustain efficiency when actions are unobservable.

In the next section, we describe a simple deterministic joint production process. There are two or more individuals, one of whom is the principal and the rest are agents, who may take part in the joint production process, but their actions in the production process are unobservable. This section also describes the preferences of the individuals and the Second Best Game. The Second Best Game is a multi stage extensive game played by the principal and all the agents to determine the set of individuals who will take part in production along with the output sharing rule they will follow. It is worth noting that the moral hazard problem in the actual production process is similar to the moral hazard problem in teams considered in section 2 of Holmstrom[11]. Hence, as mentioned later on, our efficiency result can also be derived using his results. However, unlike the present paper, Holmstrom[11] focuses attention on the issue of mitigating moral hazard in a joint production process and does not look at the problem faced by a principal who can actively participate in the production process[2].

Section 3 looks at the First Best situation and describes the appropriate notion of efficiency. We derive some optimal outcomes of the Second Best Game that involve linear output sharing rules in section 4. Section 4 also looks at the issue of mitigation of moral hazard and sustainability of efficiency. In section 5, we show that there are no significant changes in our results when there is limited liability. The uniqueness of the subgame perfect equilibrium utility tuple in the Second Best Game is established in section 6. Section 7 illustrates most of our findings in a simple example. Why our results are robust to more general utility functions is briefly discussed in section 8. Conclusions are given in section 9.

2 Production and Preferences

There are N (≥ 2) individuals who can participate in a joint production process. Whenever an individual, indexed i, participates in the production process, she takes an unobservable and/or unverifiable action $a_i \in A_i = \Re_+$. For each individual i, $c_i : A_i \rightarrow \Re_+$ is the cost function that specifies the cost she incures from her action when she participates in the production process. All inputs other than the actions of the individuals are assumed to be observable, and hence, suppressed in the specification of the model. In the production process, the actions of the individuals determine a joint monetary outcome. This production process, assumed to be deterministic, is represented by a function, $f : A \rightarrow \Re_+$, where $A = \prod_{i=1}^{N} A_i$.

For each individual i, her preference relation over money-action pairs is represented by a quasilinear utility function, $U_i : \Re \times A_i \rightarrow \Re$, which is of the form $U_i(m_i, a_i) = m_i - c_i(a_i)$ for any $(m_i, a_i) \in \Re \times A_i$. Because we consider the case in which it is possible for individuals to get negative payments, note that the utility function U_i is defined even for pairs with negative amounts of money. Later on, we discuss how the results are qualitatively affected if we abandon the quasilinear form of the utility functions.

We use the following standard notations: A_{-i} is the Cartesian product of A_j over all j not equal to i; $a = (a_1, ..., a_N) \in A$; $a_{-i} = (a_1, ..., a_{i-1}, a_{i+1}, ..., a_N) \in A_{-i}$; and $a = (a_{-i}, a_i)$.

Throughout, we assume that the production function f, the cost functions c_i, and the utility functions U_i are common knowledge. In addition, we maintain the following assumptions about the functions c_i and f :

[2] Although the outsider who administers the "budget-breaking" incentive scheme in Holmstrom's proposed solution to the moral hazard problem is often interpreted as the principal, this outsider, unlike the principal in the current paper, is precluded from taking part in production.

A1. *For each individual i, c_i is continuously differentiable, strictly increasing and strictly convex on A_i; $c_i(0) = 0$; $c_i'(0) = 0$ and $\lim_{a_i \to \infty} c_i'(a_i) = \infty$, where c_i' denotes the derivative of c_i.*

A2. *f is continuously differentiable, strictly increasing and concave on A; $f(0, ..., 0) = 0$; for each i and each $a_{-i} \in A_{-i}$, $\lim_{a_i \to 0} f_i(a_{-i}, a_i) > 0$ and $\lim_{a_i \to \infty} f_i(a_{-i}, a_i) < \infty$, where f_i is the partial derivative of f with respect to the action of individual i.*

Assumption **A1** is standard and needs no explaination. In assumption **A2**, the smoothness, monotonicity and curvature properties of the production function, and the requirement that the output be zero when everyone takes zero action are standard. Also, the limiting behaviour of the marginal product of any individual's action when it approaches infinity is standard. However, our stipulation in assumption **A2** about the limiting behaviour of the marginal product of an individual's action when it approaches zero is not so standard, because it says that the marginal product of an individual's action in a neighbourhood of zero is positive even when every other individual takes zero action. This means that, no matter what the actions of the other individuals are, the total output is strictly increasing in the action of each individual. Hence, nobody is essential for production, as total output is equal to zero only when every individual takes zero action. This particular property of the production function along with the quasilinearity of the utilities are exploited in the derivation of the results on uniqueness of equilibrium.

We treat individual N as the ***residual claimant*** in the following sense– (i) like the principal in the standard agency models, individual N is the only one who can design and propose output sharing rules; and (ii) whether she takes part in production or not, individual N always keeps that part of the output left after making payments to the other participating individuals[3]. Unlike individual N, each of the first $N-1$ individuals receives a payment only if she takes part in production. Thus, throughout the remainder of this paper, we call individual N– the *principal,* and the first $N-1$ individuals– the *agents.* However, it must be noted that, unlike the principal in standard agency models, individual N can choose to participate in the production process.

If the actions taken by the agents in the production process are unobservable and/or unverifiable, then the principal cannot make the payment to any agent depend on that agent's action. Hence, the payment to an agent for participating in the production process can depend only on the observable and/or verifiable total output of the production process. Thus, in general, a *payment function* for an agent is a real valued function defined on $\Re_+$, the set of all possible output levels. However, in this paper we shall require the payment functions to satisfy a regularity condition. This condition essentially says that the curve of the payment function of each agent has only a finite number of jumps and kinks. Thus, the payment function of each agent must be drawn from the set

$$S = \left\{ s : \Re_+ \to \Re \;\middle|\; \begin{array}{l} \text{(i)} \;\; s \text{ is piecewise continuous on } \Re_+; \text{ and} \\ \text{(ii) if } s \text{ is continuous on } (x^L, x^H) \subset \Re_+, \text{ then it is} \\ \qquad \text{piecewise continuously differentiable on } (x^L, x^H) \end{array} \right\}.$$

Although a technical restriction, requiring the payment functions of the agents be drawn from the class S seems quite reasonable, because most of the observed payment functions in joint production processes belong to this class of payment functions.

Let $\mathcal{N}$ be the set of all subsets of $\{1, ..., N\}$. So each member of $\mathcal{N}$ is a subset of the set of N individuals. We call each member of $\mathcal{N}$, a *team*, and denote them by T, $\hat{T}$, $\bar{T}$, etc. Given a team T and an individual i, we

[3] It is implicitly assumed that this residual claimant role of individual N is a consequence of reasons exogenous to the specification of the model. As mentioned in the introduction, one such good reason may be individual N owning certain tangible inputs like technology, equipment, capital, *etc.*

denote the set of all individuals in T other than i by T_{-i}; *i.e.* $T_{-i} = T - \{i\}$. Thus, $T = T_{-i}$ if and only if $i \notin T$.

Given $T \in \mathcal{N}$, let A^T be the Cartesian product of A_i over all i in T, and $a_T = (a_i)_{i \in T} \in A^T$. Also, given any $T \in \mathcal{N}$, let $f|T : A^T \to \Re_+$ be the restriction of f to A^T in the following sense– for each $a_T \in A^T$, $f|T(a_T) = f(a')$, where $a' = (a'_1, ..., a'_N) \in A$ is such that $a'_i = a_i$ if $i \in T$ and $a'_i = 0$ if $i \notin T$. For each $T \in \mathcal{N}$ and each $i \in T$, we use $f_i|T$ to denote the partial derivative of $f|T$ with respect to the action of individual i.

As production can take place with the participation of any subset of individuals, it is clear that, when only the members of some $T \in \mathcal{N}$ take part in production, the relevant production process is $f|T$.

Given any $T \in \mathcal{N}$, if only the members of T take part in production, then an *output sharing rule* for T is a tuple of payment functions, $(s_i)_{i \in T_{-N}}$, where $s_i \in S$ is the payment function of agent $i \in T_{-N}$. Obviously, for each output level $x \in \Re_+$, whether the principal belongs to T or not, she gets the residual $x - \sum_{i \in T_{-N}} s_i(x)$.

Each individual has an outside option which she can exercise instead of taking part in the production process. The utility of individual i from her outside option is equal to $u_i \geq 0$. So, when agent i exercises her outside option, she automatically gets zero payment from the principal and her utility is equal to u_i. However, as the principal is the residual claimant, she still gets her residual in addition to u_N even when she exercises her outside option.

Because the action taken by any individual in the production process is neither observable nor verifiable, once the team which will take part in production along with the output sharing rule become common knowledge, the members of the team actually play a noncooperative game in the production process conditional on the common knowledge output sharing rule. Suppose it becomes common knowledge that a team T will take part in production and the output sharing rule will be $(s_i)_{i \in T_{-N}}$. Then the strategies and payoffs of the players in the ensuing *noncooperative game of production* (NGP), which we denote by $\{T, (s_i)_{i \in T_{-N}}\}$, are as follows: (i) a strategy of player $j \in T$ is an action $a_j \in A_j$; and (ii) when the actions taken by the players in T are $a_T \in A^T$, the payoff of player $j \in T$ is $s_j(f|T(a_T)) - c_j(a_j)$ if j is an agent (*i.e.* if $j \neq N$) and $f|T(a_T) - \sum_{i \in T_{-N}} s_i(f|T(a_T)) - c_N(a_N)$ if j is the principal (*i.e.* if $j = N$).

Given any NGP $\{T, (s_i)_{T_{-N}}\}$ such that $N \notin T$, $a_T \in A^T$ is a Nash equilibrium of this NGP if and only if

$$a_i \in \text{argmax}_{a'_i \in A_i} [s_i(f|T(a_{T_{-i}}, a'_i)) - c_i(a'_i)] \quad \forall\, i \in T.$$

Similarly, given any NGP $\{T, (s_i)_{T_{-N}}\}$ such that $N \in T$, $a_T \in A^T$ is a Nash equilibrium of this NGP if and only if

$$\begin{aligned} a_i &\in \text{argmax}_{a'_i \in A_i} [s_i(f|T(a_{T_{-i}}, a'_i)) - c_i(a'_i)] \quad \forall\, i \in T_{-N}; \text{ and} \\ a_N &\in \text{argmax}_{a'_N \in A_N} [f|T(a_{T_{-N}}, a'_N) - \sum_{i \in T_{-N}} s_i(f|T(a_{T_{-N}}, a'_N)) - c_N(a'_N)]. \end{aligned}$$

We denote the set of all Nash equilibria of each NGP, $\{T, (s_i)_{i \in T_{-N}}\}$, by $NE(\{T, (s_i)_{i \in T_{-N}}\})$.

For each team in $\mathcal{N}$, it is clear that there are countless number of possible output sharing rules. This means that, as production can take place with the participation of any one team in $\mathcal{N}$, there are countless number of possible team and output sharing rule combinations according to which production can take place. So we need a procedure that determines a single team and output sharing rule combination according to which actual production takes place.

Suppose the principal deals secretly with different subgroups of agents and ultimately arrives at a single team T and a corresponding output sharing rule $(s_i)_{i \in T_{-N}}$ according to which actual production takes place.

Then it is very likely that the output sharing rule $(s_i)_{i\in T_{-N}}$ is not common knowledge among the agents in T. So the strategic behaviour of each agent in T in the production process depends on her belief about the output sharing rule, her beliefs about the beliefs of the others in T and so on. Clearly, the beliefs of each agent in T depend on all the information that she has, for example, her own payment function and may be that of some of the other agents in T. However, given all the information that is available to an agent, the manner in which she uses them to form her beliefs is quite complicated to model and well beyond the scope of this paper. Therefore, we assume that the principal does not deal secretly with any subgroup of agents and whatever she proposes becomes common knowledge immediately.

As we assume that no one can be forced to participate in production, even though the principal, as the residual claimant, is the only individual who can propose any team and output sharing rule combination, an agreement must be reached on a single team and output sharing rule combination according to which actual production takes place. However, handling the complex strategic issues involved when individuals are allowed to collude with one another in trying to reach an agreement on a single team and output sharing rule combination are well beyond the scope of this paper. Moreover, it is seldom easy to justify the credibility of commitment of any member of a coalition to the coalition. Therefore, we assume that the individuals behave noncooperatively when trying to reach an agreement on a single team and output sharing rule combination.

Thus, we use a very simple multi-stage procedure to determine a single team and output sharing rule combination according to which production takes place. This multi-stage procedure, which we call the *Second Best Game* (SBG), is described as follows:

Stage I: In the first stage, the principal announces a NGP which becomes common knowledge immediately.

Stage II: In the second stage, each agent who is a player of the NGP announced by the principal must announce whether she agrees to play this NGP or not. These announcements by the agents are made sequentially, so that, the announcement of an agent becomes common knowledge before the announcement of any subsequent agent. If any agent who is a player of the NGP announced by the principal announces a disagreement, then the procedure terminates at this point and everyone exercise their respective outside options. Of course, those individuals who are not players of the NGP announced by the principal automatically exercise their outside options.

Stage III: This stage is reached only if every agent who is a player of the NGP announced by the principal announced an agreement in the second stage. Once the third stage is reached, the NGP announced by the principal is played in the production process.

Obviously, if the third stage is not reached, then the payoff of each individual i in the SBG is her outside option utility, u_i. However, if the third stage is reached, then the payoffs in the SBG are given as follows: (i) each agent who is not a player of the NGP played in the third stage gets her outside option utility; (ii) each agent who is a player of the NGP played in the third stage gets her payoff from this NGP; (iii) if the principal is a player of the NGP played in the third stage, then she gets her payoff from this NGP; and (iv) if the principal is not a player of the NGP played in the third stage, then she gets u_N (her outside option utility) plus the output left after distributing the payments to the players of the NGP played in the third stage.

Given $T \in \mathcal{N}$ and an output sharing rule $(s_i)_{i\in T_{-N}}$ such that $s_i \in S$ for each $i \in T_{-N}$, there is no gaurantee that the NGP $\{T, (s_i)_{i\in T_{-N}}\}$ has a Nash equilibrium. However, as we want to focus only on subgame perfect equilibria of the SBG, we cannot allow the principal to propose NGPs that do not have any Nash equilibrium. Hence, we impose the restriction that the NGP announced by the principal in the first stage of the SBG be drawn from the set

$$\mathcal{G} = \left\{ \{T, (s_i)_{i \in T_{-N}}\} \;\middle|\; \begin{array}{l} \text{(i) } T \in \mathcal{N}; \text{ (ii) } s_i \in S \; \forall \; i \in T_{-N}; \text{ and} \\ \text{(iii) } NE(\{T, (s_i)_{i \in T_{-N}}\}) \neq \emptyset \end{array} \right\}.$$

3 First Best

To understand the efficiency properties of the model it is necessary to know the meaning of efficiency in the present context. So, as in most standard moral hazard models, we look at the *First Best* (FB) situation in order to find an appropriate notion of efficiency for our model.

The FB situation is the hypothetical situation in which the action of each individual in the production process is observable, and hence, the principal pays each agent who takes part in production according to her action. So, in the FB situation the principal can dictate the action of each agent in the following sense– if the principal wants an agent to participate in production and take a particular level of action, then she can solicit the desired action voluntarily from the agent with a sufficient payment for that action and zero payment for any other action. Therefore, when only the members of a team $T \in \mathcal{N}$ take part in production in the FB situation, the principal chooses a tuple of payment-action pairs for all the agents in T and an action for herself if she belongs to T to maximize her utility subject to the condition that each agent in T gets at least as much utility as from her outside option. We denote this maximum utility of the principal by $B(T)$.

Suppose $T \in \mathcal{N}$ is the set of individuals who take part in production in the FB situation. Then $N \notin T$ means that the principal exercises her outside option and also gets the residual. So, if $N \notin T$, then $B(T)$ is given by

$$B(T) = u_N + \max_{(a_T, m_T)} \left\{ f|T(a_T) - \sum_{i \in T} m_i \;\middle|\; \begin{array}{l} \text{(i) } m_i - c_i(a_i) \geq u_i \; \forall \; i \in T; \\ \text{(ii) } a_T \in A^T; \text{ and} \\ \text{(iii) } m_T = (m_i)_{i \in T} \in \Re^{|T|} \end{array} \right\}.$$

On the otherhand, the principal can belong to T only if she does not exercise her outside option. Hence, if T is such that $N \in T$, then $B(T)$ is defined by

$$B(T) = \max_{(a_T, m_{T_{-N}})} \left\{ f|T(a_T) - \sum_{i \in T_{-N}} m_i - c_N(a_N) \;\middle|\; \begin{array}{l} \text{(i) } m_i - c_i(a_i) \geq u_i \; \forall \; i \in T_{-N}; \\ \text{(ii) } a_T \in A^T; \text{ and} \\ \text{(iii) } m_{T_{-N}} = (m_i)_{i \in T_{-N}} \in \Re^{|T_{-N}|} \end{array} \right\}.$$

In the definition of $B(T)$ in either case, the constraints, $m_i - c_i(a_i) \geq u_i \; \forall \; i \in T_{-N}$, ensure that each participating agent is no worse off than exercising her outside option. These constraints are obviously necessary for the participating agents to be willing participants. Also, we must point out in passing that $B(T)$ is well defined for any $T \in \mathcal{N}$, because, using assumptions **A1** and **A2**, it can be easily shown that the maximization problem in the definition of $B(T)$ in either case has a solution.

Suppose the principal chooses to exercise her outside option in the FB situation. Clearly, if $B(T) \leq u_N$ for every $T \in \mathcal{N}$ such that $N \notin T$, then the highest utility she can get is u_N. On the otherhand, if there exists $T \in \mathcal{N}$ such that $N \notin T$ and $B(T) > u_N$, then the highest utility she can get is the maximum of $B(T)$ over all $T \in \mathcal{N}$ such that $N \notin T$. Thus, if the principal chooses to exercise her outside option in the FB situation, then the highest utility she can get, which we denote by u^F_{N-}, is given by

$$u^F_{N-} = \begin{cases} u_N & \text{if } B(T) \leq u_N \; \forall \; T \in \mathcal{N} \text{ such that } N \notin T \\ \max_{T \in \mathcal{N}} \{B(T) | \; N \notin T\} & \text{otherwise.} \end{cases}$$

perspective) in $\hat{\Omega}_{-N} \subseteq \hat{\Omega}$, which is the set of all outcomes in $\hat{\Omega}$ where the principal does not participate in production; *i.e.*

$$\hat{\Omega}_{-N} = \left\{(\{T, (s_i)_{i \in T_{-N}}\}, a_T) \in \hat{\Omega} \mid N \notin T\right\}.$$

Let T^F be a team in $\mathcal{N}$ such that $N \notin T^F$, and $B(T^F) = \max_{T \in \mathcal{N}}\{B(T) \mid N \notin T\}$. Also, let $(m^F_{T^F}, a^F_{T^F}) \in \Re^{|T^F|} \times A^{T^F}$ be payment and action tuples such that $m_i^F - c_i(a_i^F) \geq u_i \; \forall \; i \in T$, and $u_N + f|T^F(a^F_{T^F}) - \sum_{i \in T^F} m_i^F = B(T^F)$. So the team T^F along with the payment and action tuples $(m^F_{T^F}, a^F_{T^F})$ give the highest utility to the principal if she does not participate in production in the FB situation. Clearly, $B(T^F) \geq \pi((\{T, (s_i)_{i\in T}\}, a_T)) \; \forall \; (\{T, (s_i)_{i \in T}\}, a_T) \in \hat{\Omega}_{-N}$. Then, according to assumptions **A3** and **A5**, $B(T^F) > u_N$ and $|T^F| \geq 2$ are necessary for any outcome in $\hat{\Omega}_{-N}$ to be an OO. Thus, unless otherwise mentioned, it must be understood that we are only looking at the case in which $B(T^F) > u_N$ and $|T^F| \geq 2$.

Because of Lemma 1, we know that the following are true:

$$(1) \qquad m_i^F - c_i(a_i^F) = u_i \quad \forall \, i \in T;$$

$$(2) \qquad u_N + f|T^F(a^F_{T^F}) - \sum_{i \in T^F} m_i^F = B(T^F); \text{ and}$$

$$(3) \qquad f_i|T(a^F_{T^F}) - c_i'(a_i^F) = 0 \quad \forall \, i \in T.$$

So we seek to construct an output sharing rule which will induce the agents in T^F to take the actions $a^F_{T^F}$ and also pay m_i^F to each agent $i \in T^F$ at the output level $f|T^F(a^F_{T^F})$. Now, (3) says that any output sharing rule which induces the actions $a^F_{T^F}$ and is smooth in a neighbourhood of the output level $f|T^F(a^F_{T^F})$ must only have payment functions that have unit slopes around a neighbourhood of the output level $f|T^F(a^F_{T^F})$. However, because of the deterministic nature of the production function f, (3) does not say anything about how the output sharing rule should behave away from a neighbourhood of the output level $f|T^F(a^F_{T^F})$. This gives sufficient freedom that allows us to construct a desired output sharing rule which is linear.

Suppose the agents in T^F play a NGP in which the payment to each agent $i \in T^F$ is equal to the total output plus the constant $u_i + c_i(a_i^F) - f|T^F(a^F_{T^F})$ for every level of output. Then the quasilinear utility functions, the strict convexity of the cost functions and the concavity of the production function imply that a tuple of actions for the agents in T^F is a Nash equilibrium if the marginal product (which is the same as the marginal benefit) is equal to the marginal cost for every agent in T^F. However, we already know from (3) that the marginal product is equal to the marginal cost for each agent in T^F at $a^F_{T^F}$. Therefore, $a^F_{T^F}$ is a Nash equilibrium. Also, it is easy to check that the utilities of each agent $i \in T^F$ and the principal at $a^F_{T^F}$ are u_i and $B(T^F)$, respectively. Furthermore, because of the quasilinear utility functions and the monotonicity, curvature and limiting marginal properties of the cost functions and the production function, $a^F_{T^F}$ is in fact the unique Nash equilibrium.

For each $i \in T^F$, let α_i^F be the constant such that

$$(4) \quad \alpha_i^F = u_i + c_i(a_i^F) - f|T^F(a^F_{T^F}) \quad \forall \, i \in T^F.$$

Now, for each agent $i \in T^F$, define the payment function s_i^F as follows:

$$(5) \quad s_i^F(x) = \alpha_i^F + x \qquad \forall \, x \in \Re_+.$$

Then, more formally, we have the following proposition.

Proposition 1 : *If assumptions* **A1** *and* **A2** *are satisfied, then:* (i) $(\{T^F, (s_i^F)_{i \in T^F}\}, a^F_{T^F}) \in \hat{\Omega}_{-N}$; (ii) $s_i^F(f|T^F(a^F_{T^F})) - c_i(a_i^F) = u_i \; \forall \; i \in T^F$; (iii) $u_N + f|T^F(a^F_{T^F}) - \sum_{i \in T^F} s_i^F(f|T^F(a^F_{T^F})) = B(T^F)$; *and* (iv) $NE(\{T^F, (s_i^F)_{i \in T^F}\}) = \{a^F_{T^F}\}$.

Proof: See Appendix A.

The intuition behind Proposition 1 is as follows. As the principal is the residual claimant, when the actions are not observable and only the individuals in T^F take part in production, the role of the principal in this paper and the role of the outsider who administers a budget-breaking incentive scheme in Holmstrom[11] are the same. So, whenever the total output deviates from $f|T^F(a^F_{T^F})$, although the principal does not know the agent(s) whose action(s) caused this deviation, she can find the entire team of agents, T^F, at fault. Hence, the principal can punish everyone in T^F for any deviation in the total output from $f|T^F(a^F_{T^F})$. In particular, the principal can make each agent $i \in T^F$ fully responsible for any deviation in the total output from $f|T^F(a^F_{T^F})$ by paying her the total output plus the constant $u_i + c_i(a_i^F) - f|T^F(a^F_{T^F})$ for every level of output.

Next, let us look at the other half of the set $\hat{\Omega}$, denoted by $\hat{\Omega}_{+N}$, which contains all those outcomes from $\hat{\Omega}$ where the principal participates in production; *i.e.*

$$\hat{\Omega}_{+N} \;=\; \Big\{(\{T,(s_i)_{i\in T_{-N}}\},a_T) \in \hat{\Omega}|\;\; N \in T\Big\}.$$

When the principal is the only player, there is only one NGP, namely, the one in which the principal keeps the entire output for herself. Although it is obvious that the principal can get $B(\{N\})$ in this NGP, because of assumption **A5**, there cannot be any OO which involves this NGP. So we only need to pay attention to those outcomes in $\hat{\Omega}_{+N}$ that have at least one agent participating in production along with the principal. Also, as we are interested in OOs, assumption **A4** allows us to ignore those outcomes in $\hat{\Omega}_{+N}$ at which some participant in the production process takes zero action.

Thus, among all the outcomes in $\hat{\Omega}_{+N}$ that have two or more individuals participating in production with everyone of them taking a positive action, we are interested only on those that are best from the principal's perspective. Therefore, as an intermediate step, for each $T \in \mathcal{N}$ such that $N \in T$ and $|T| \geq 2$, we need to look at the following maximization problem:

$$(P_T) \qquad \max_{a_T,(s_i)_{i\in T_{-N}}} \Big[f|T(a_T) - \textstyle\sum_{i\in T_{-N}} s_i(f|T(a_T)) - c_N(a_N)\Big]$$

subject to:

(C1) $\quad a_i > 0 \quad \forall\, i \in T$; and

(C2) $\quad (\{T,(s_i)_{i\in T_{-N}}\},a_T) \;\in\; \hat{\Omega}_{+N}$.

As discussed above, constraint (C1) requires a positive action for every individual in T. Constraint (C2) obviously follows from the fact that SPE outcomes have to be in the set $\hat{\Omega}$.

Clearly, constraint (C2) of problem (P_T) involves maximization problems of the players in T. So we use a standard method, commonly known as the ***first order approach***, to solve problem (P_T). As the first order approach uses only the necessary conditions of the optimization problems involved in the constraint, sometimes the solution(s) obtained by using this approach may not be solution(s) of the original problem. However, we need not worry about such a possibility in the present case, because the solutions we derive by using the first order approach are indeed solutions of problem (P_T).

Appendix A proves a technical lemma that enables us to use the first order approach. Given $T \in \mathcal{N}$ such that $N \in T$ and $|T| \geq 2$, if a NGP and an action tuple corresponding to T satisfy (C1) and (C2), then this lemma asserts that the curve of the payment function of each agent in T is smooth at the output level corresponding to the given action tuple.

Lemma 3 *: Suppose assumptions* **A1** *and* **A2** *are satisfied, and $T \in \mathcal{N}$ is such that $N \in T$ and $|T| \geq 2$. If $(\{T,(s_i)_{i\in T_{-N}}\},a_T)$ satisfies* (C1) *and* (C2)*, then s_i is differentiable at $f|T(a_T)$ for each $i \in T_{-N}$.*

Proof: See Appendix A.

Suppose $(\{T,(s_i)_{i\in T_{-N}}\},a_T) \in \hat{\Omega}_{+N}$ is such that $|T| \geq 2$ and $a_i > 0 \ \forall\ i \in T$. Because of the quasilinear utility functions, it is clear that the marginal benefit of each individual in T is equal to her marginal cost at a_T. Also, because of Lemma 3, we know that, for each agent $i \in T_{-N}$, her marginal benefit at a_T can be written as the product of the slope of her payment function s_i at $f|T(a_T)$ and her marginal product at a_T. Hence, for each agent $i \in T_{-N}$, the slope of s_i at $f|T(a_T)$ must be equal to the ratio of her marginal cost and marginal product at a_T. As the principal's residual for any level of output $x \in \Re_+$ is $x - \sum_{i\in T_{-N}} s_i(x)$, Lemma 3 also implies that the marginal benefit of the principal at a_T is equal to her marginal product at a_T times one minus the sum of the slopes of the payment functions of all the agents in T at $f|T(a_T)$. Therefore, one minus the sum of the slopes of the payment functions of all the agents in T at $f|T(a_T)$ must be equal to the ratio of the principal's marginal cost and marginal product at a_T. But we already know that the slope of the payment function of agent $i \in T_{-N}$ at $f|T(a_T)$ is equal to the ratio of her marginal cost and marginal product at a_T. So the marginal cost to marginal product ratios of all the individuals in T at a_T must add up to one. More formally, we have the following lemma.

Lemma 4 *: Suppose assumptions* **A1** *and* **A2** *are satisfied, and $T \in \mathcal{N}$ such that $N \in T$ and $|T| \geq 2$. If $(\{T,(s_i)_{i\in T_{-N}}\},a_T)$ satisfies* (C1) *and* (C2), *then* $\sum_{i\in T}[c_i'(a_i)/f_i|T(a_T)] = 1$.

Proof: See Appendix A.

Given $T \in \mathcal{N}$ such that $N \in T$ and $|T| \geq 2$, suppose we replace (C1) and (C2) by the marginal condition in Lemma 4. Also, for each $i \in T_{-N}$, suppose we replace $s_i(f|T(a_T))$ by $u_i + c_i(a_i)$ in the objective function of problem (P_T). Then we get the following new maximization problem:

$$(P_T') \qquad \max_{a_T\in A^T} \left[f|T(a_T) - \textstyle\sum_{i\in T} c_i(a_i) - \sum_{i\in T_{-N}} u_i\right]$$

subject to:

$$\text{(C3)} \qquad \textstyle\sum_{i\in T}[c_i'(a_i)/f_i|T(a_T)] \quad = \quad 1.$$

Note that the actions of the individuals in T are no longer required to be positive in problem (P_T'). Also, unlike problem (P_T), the only choice variables in problem (P_T') are the actions of the individuals in T. The intuitive logic behind the transformation of problem (P_T), which involves payment functions, into problem (P_T'), which does not involve any payment function, is as follows. Because of the deterministic nature of the production function f, constraints (C1) and (C2) only tell how the output sharing rule $(s_i)_{i\in T_{-N}}$ must behave around a neighbourhood of the output level $f|T(a_T)$ and not else where. Then, as the payments to the agents and the principal's residual has to add up to the total output, this local condition implied by (C1) and (C2) translates into (C3) and eliminates the payment functions.

Lemma 5 *: Suppose assumptions* **A1** *and* **A2** *hold, and $T \in \mathcal{N}$ such that $N \in T$ and $|T| \geq 2$. Then problem (P_T') has a solution.*

Proof: See Appendix A.

Given any $T \in \mathcal{N}$ such that $N \in T$ and $|T| \geq 2$, the value of the objective function of problem (P_T') at a solution is denoted by $V(T)$. So, if a_T is a solution of problem (P_T'), then $f|T(a_T) - \sum_{i\in T} c_i(a_i) - \sum_{i\in T_{-N}} u_i = V(T)$.

Remember that, for any combination of $(s_i)_{i\in T_{-N}}$ and a_T which is feasible for problem (P_T), the individual rationality conditions of the agents in T, $s_i(f|T(a_T)) - c_i(a_i) \geq u_i \ \forall\ i \in T_{-N}$, are included in constraint (C2).

So the value of the objective function of problem (P_T) at $(a_T, (s_i)_{i\in T_{-N}})$ cannot be greater than the value of the objective function of problem (P'_T) at a_T. But, because of Lemma 4, the action tuple at any feasible point of problem (P_T) is also feasible for problem (P'_T). Therefore, we cannot find a feasible point of problem (P_T) at which the value of its objective function is greater than the value of the objective function of problem (P'_T) at a solution.

Lemma 6 : *Suppose assumptions* **A1** *and* **A2** *are satisfied, and* $T \in \mathcal{N}$ *such that* $N \in T$ *and* $|T| \geq 2$. *If* $(a_T, (s_i)_{i\in T_{-N}})$ *is feasible for problem* (P_T), *then* $V(T) \geq f|T(a_T) - \sum_{i\in T_{-N}} s_i(f|T(a_T)) - c_N(a_N)$.

Proof: See Appendix A.

Let T^* maximize $V(T)$ over all $T \in \mathcal{N}$ such that $N \in T$ and $|T| \geq 2$; *i.e.*

$$T^* \in \operatorname{argmax}_{T\in\mathcal{N}} \{V(T) \mid N \in T \text{ and } |T| \geq 2\}.$$

The existence of T^* follows from Lemma 5 and the finiteness of the number of teams that have the principal and at least one agent. Also, let $a^*_{T^*} \in A^{T^*}$ be a solution of problem (P'_{T^*}). Then we have

$$(6) \qquad \sum_{i\in T^*} [c'_i(a^*_i)/f_i|T^*(a^*_{T^*})] = 1; \quad \text{and}$$

$$(7) \qquad f|T^*(a^*_{T^*}) - \sum_{i\in T^*} c_i(a^*_i) - \sum_{i\in T^*_{-N}} u_i = V(T^*).$$

Thus, we want an output sharing rule which will induce the action tuple $a^*_{T^*}$ and pay $u_i + c_i(a^*_i)$ to each agent $i \in T^*_{-N}$ at the output level $f|T^*(a^*_{T^*})$. However, as mentioned above, because of the absence of uncertainty in the production process, we know that we have some freedom in choosing the behaviour of the desired output sharing rule away from the output level $f|T^*(a^*_{T^*})$. Below, we show that this freedom is indeed enough for us construct a linear output sharing rule.

For each $i \in T^*_{-N}$, let the two constants β^*_i and α^*_i be as follows:

$$(8) \qquad \beta^*_i = c'_i(a^*_i)/f_i|T^*(a^*_{T^*}); \quad \text{and}$$

$$(9) \qquad \alpha^*_i = u_i + c_i(a^*_i) - \beta^*_i f|T^*(a^*_{T^*}).$$

Then, for each $i \in T^*_{-N}$, let s^*_i be the linear payment function whose slope is β^*_i and intercept is α^*_i; *i.e.*

$$(10) \qquad s^*_i(x) = \alpha^*_i + \beta^*_i x \quad \forall\, x \in \Re_+.$$

Clearly, for each $i \in T^*_{-N}$, the slope of s^*_i, β^*_i, is nonnegative and equal to zero only if $a^*_i = 0$, which we have not yet ruled out. Also, $(\alpha^*_i)_{i\in T^*_{-N}}$, the intercepts of the payment functions in $(s^*_i)_{i\in T^*_{-N}}$, are set in such a way that, if $(s^*_i)_{i\in T^*_{-N}}$ is the output sharing rule and $a^*_{T^*}$ is the action tuple taken by the individuals in T^*, then the utilities of agent $i \in T^*_{-N}$ and the principal are u_i and $V(T^*)$, respectively. So Lemma 6 implies that $(a^*_{T^*}, (s^*_i)_{i\in T^*_{-N}})$ is indeed a solution of problem (P_{T^*}) if feasible.

Suppose agent $i \in T^*_{-N}$ is paid according to s^*_i and the actions of the other individuals in T^* are fixed at $a^*_{T^*_{-i}}$. Then, because of her quasilinear utility function, the utility of agent i as a function only of her own action, $a_i \in A_i$, can be separated into the benefit function, $\alpha^*_i + \beta^*_i f|T^*(a^*_{T^*_{-i}}, a_i)$, and the cost function, $c_i(a_i)$. We know that the cost function is strictly increasing. However, the benefit function is just the constant α^*_i if β^*_i is equal to zero, and strictly increasing if β^*_i is positive. So it is obvious that, if $\beta^*_i = 0$, which can happen only if $a^*_i = 0$, then the best action for agent i is a^*_i $(= 0)$. On the otherhand, if $a^*_i > 0$, then β^*_i is positive and the benefit function is strictly increasing, but there is a trade-off between the increase in the benefit and the increase in the cost as the action of agent i increases. Then it makes sense for agent i to choose a_i^* if $a_i^* > 0$, because her marginal benefit is equal to her marginal cost at a_i^*.

When the agents in T^* are paid according to the output sharing rule $(s_i^*)_{i \in T^*_{-N}}$, the residual function of the principal, $-\sum_{i \in T^*_{-N}} \alpha_i^* + (1 - \sum_{i \in T^*_{-N}} \beta_i^*)x$, is also linear in the output $x \in \Re_+$. Because of (6) and (8), $(1 - \sum_{i \in T^*_{-N}} \beta_i^*)$, the slope of the residual function, is equal to $c'_N(a_N^*)/f_N|T^*(a^*_{T^*})$. Then, using a similar intuition as in the case of the agents, we can say that a_N^* is the best action for the principal if the output sharing rule is $(s_i^*)_{i \in T^*_{-N}}$ and the actions of the agents in T^* are fixed at $a^*_{T^*_{-N}}$.

Proposition 2 : *If assumptions* **A1-A5** *hold, and* $V(T^*) > B(T^F)$, *then:* (i) $(a^*_{T^*}, (s_i^*)_{i \in T^*_{-N}})$ *is a solution of problem* (P_{T^*}); (ii) $s_i^*(f|T^*(a^*_{T^*})) - c_i(a_i^*) = u_i \; \forall \, i \in T^*_{-N}$; (iii) $f|T^*(a^*_{T^*}) - \sum_{i \in T^*_{-N}} s_i^*(f|T^*(a^*_{T^*})) - c_N(a_N^*) = V(T^*)$; *and* (iv) $NE(\{T^*, (s_i^*)_{i \in T^*_{-N}}\}) = \{a_{T^*}\}$.

Proof: See Appendix A.

The condition, $V(T^*) > B(T^F)$, plays a crucial role in Proposition 2. Whenever it holds, because of assumptions **A4** and **A5**, the principal's utility at an OO cannot exceed $V(T^*)$. So, once we establish $(\{T^*, (s_i^*)_{i \in T^*_{-N}}\}, a^*_{T^*}) \in \hat{\Omega}$ and $\pi((\{T^*, (s_i^*)_{i \in T^*_{-N}}\}, a^*_{T^*})) = V(T^*)$, $a_i^* > 0$ for every $i \in T^*$ follows from assumption **A4**.

Also, note that $a^*_{T^*}$ is the unique Nash equilibrium of $\{T^*, (s_i^*)_{i \in T^*_{-N}}\}$ according to (iv) in Proposition 2. This result follows from the quasilinearity of the utility functions, the concavity of the production function and the strict convexity of the cost functions.

The most obvious but important message of Proposition 2 is that, if it is better for the principal to participate in production, then she need not look for any output sharing rule that is more sophisticated than those in the class of simple linear output sharing rules.

In contrast to the output sharing rule in Proposition 1, there is pure sharing in the output sharing rule in Proposition 2 in the sense that– every participating individual gets a constant (which may be negative) plus a positive proportion of the total output. This can be roughly interpreted as follows. If the principal is better off participating in production, then the principal can only get worse off with an output sharing rule which punishes only a particular proper subset of the set of participating individuals for every deviation in the total output from the optimal output.

If the principal is better off participating in production, then, because of (iii) of Lemma 2, the only way she can get the FB utility is if the actions taken in the production process are such that the marginal product is equal to the marginal cost for each participating agent. But, Proposition 2 says that this cannot happen, because the slopes of the payment functions and the residual function of the principal in the output sharing rule in Proposition 2 are all less than one. Thus, another important implication of Proposition 2 is that, if it is better for the principal to participate in production, then she cannot completely mitigate the moral hazard problem; *i.e.* $V(T^*) < B(T^*)$. This highlights the presence of an inherent conflict between two things, namely, the principal's role as a residual claimant and her role as a free rider whenever she participates in production.

(iv) of Proposition 2 also has an important implication. As we shall demonstrate later, this result along with the quasilinearity of the utility functions can be used to show the uniqueness of the SPE utility tuple.

Now, what we originally set out to show, namely, there is some OO in Ω^* where the output sharing rule is linear, is a rather obvious corollary of Propositions 1 and 2.

Corollary: *Suppose assumptions* **A1-A5** *hold. If* $B(T^F) \geq V(T^*)$, *then* $(\{T^F, (s_i^F)_{i \in T^F}\}, a^F_{T^F}) \in \Omega^*$. *If* $V(T^*) > B(T^F)$, *then* $(\{T^*, (s_i^*)_{i \in T^*_{-N}}\}, a^*_{T^*}) \in \Omega^*$.

From our derivations up to this point, we can naturally draw the following conclusions about the principal's participation decision: (i) if $B(T^F) > V(T^*)$, then the principal will choose not to participate in production; (ii) if $V(T^*) > B(T^F)$, then the principal will choose to participate in production; and (iii) if $B(T^F) = V(T^*)$, then the principal may or may not choose to participate in production.

Let us now look at the principal's ability to sustain efficiency. Obviously, the only way the principal can sustain efficiency is if her utility at every OO is equal to u_N^F. This means that the principal can sustain efficiency if $B(T^F) = u_N^F$, because we know that in this case the principal's utility at any OO is equal to $B(T^F)$. So the question remains, what if $u_N^F > B(T^F)$? If $u_N^F > B(T^F)$, then it is clear that the only way the principal can sustain efficiency is if $u_N^F = V(T^*)$. We know that u_N^F is at least as large as $B(T^*)$. But we also argued above that $B(T^*)$ is greater than $V(T^*)$. Therefore, $u_N^F > V(T^*)$ always holds. So the principal cannot sustain efficiency if $u_N^F > B(T^F)$. Thus, we can claim the following.

Claim 2 : *Suppose assumptions* **A1-A5** *hold. Then the principal can sustain efficiency if and only if* $u_N^F = B(T^F)$.

As the principal does not belong to T^F, the condition $u_N^F = B(T^F)$ means that the principal does not participate in production in the FB situation to obtain the maximum utility. So Claim 2 can also be put in a slightly different way as follows. To sustain efficiency it is necessary and sufficient that the principal obtain the maximum utility in the FB situation without her participation in production.

As the principal plays the role of a residual claimant, if she takes part in production, we get a joint production process to which Theorem 1 of Holmstrom[11] is applicable. On the otherhand, if the principal does not take part in production, she is just like the outsider in Holmstrom's solution to the moral hazard problem, who administers "budget-breaking" incentive schemes. Therefore, Claim 2 can also be viewed as an implication of the results in Holmstrom[11].

The intuition behind Claim 2 is as follows. When the actions are not observable, every individual who takes part in production (including the principal) has an incentive to free ride in the production process. So, when only the members of some team $T \in \mathcal{N}$ take part in production, to get $B(T)$, the FB situation utility, the principal must design an output sharing rule which has sufficient punishments for everyone in T for any deviation of the total output from the FB situation output level corresponding to $B(T)$. But, whenever the principal punishes every agent in T, as the residual claimant she can only reward herself, which means that the principal cannot punish everyone in T if she herself is a member of T. So, whenever the principal takes part in production along with a group of agent(s), there is bound to be an inherent conflict between her residual claimant role and her incentive to free ride in the production process. Therefore, when the actions are not observable and only the members of some team T take part in production, the principal can get $B(T)$ only if she does not belong to T.

5 Limited Liability

So far we have allowed output sharing rules that can award sufficiently large negative payments to some agents or the principal. However, such output sharing rules may no longer be feasible if individuals have limited liabilities. Thus, in this section we look at the case where individuals do have limited liabilities. In particular, we impose an extreme form of limited liability constraint, namely, no one, including the principal, can commit to any amount of negative payment. So the NGP announced by the principal in the first stage of the SBG must be drawn from the following set:

$$\mathcal{G}_+ = \left\{ \{T, (s_i)_{i\in T_{-N}}\} \in \mathcal{G} \mid \begin{array}{l} \text{(i) } s_i(x) \geq 0 \ \forall\ x \in \Re_+ \text{ and each } i \in T_{-N}\text{; and} \\ \text{(ii) } x - \sum_{i\in T_{-N}} s_i(x) \geq 0 \ \forall\ x \in \Re_+ \end{array} \right\}.$$

The question we ask is, is there any OO $(\{T, (s_i)_{i\in T_{-N}}\}, a_T) \in \Omega^*$ such that $\{T, (s_i)_{i\in T_{-N}}\}$ belongs to $\mathcal{G}_+$? The answer is yes. In fact, we modify the linear output sharing rule in the appropriate OO derived in the previous section in such a way that the limited liability constraint is met and the modified output sharing rule along with the original action tuple remains an OO. This modification is carried out in such a way that the payment functions of the agents (and the residual function of the principal if $V(T^*) > B(T^F)$) are continuous, piecewise linear and nondecreasing.

The reason that allows us to perform our modifications is quite obvious. Suppose $B(T^F) \geq V(T^*)$ ($V(T^*) > B(T^F)$). Then, as long as we keep the output sharing rule $(s_i^F)_{i\in T^F}$ ($(s_i^*)_{i\in T^*_{-N}}$) intact on an appropriate range of output around $f|T^F(a^F_{T^F})$ ($f|T^*(a^*_{T^*})$), the absence of uncertainty in the production process provides enough freedom that allows us to change $(s_i^F)_{i\in T^F}$ ($(s_i^*)_{i\in T^*_{-N}}$) quite arbitrarily else where such that the action tuple $a^F_{T^F}$ ($a^*_{T^*}$) is still induced.

Let us first consider the case $B(T^F) \geq V(T^*)$. Then it can be easily checked that, for each $i \in T^F$, $\alpha_i^F < 0$ and the payment function s_i^F awards negative payments to agent i only at output levels below $-\alpha_i^F$. Also, the principal's residual, $x - \sum_{i\in T^F} s_i^F(x)$, becomes negative only beyond a certain output level greater than $f|T^F(a^F_{T^F})$. So, for each $i \in T^F$, we can modify s_i^F in such a way that there are no changes between the output levels $-\alpha_i^F$ and $f|T^F(a^F_{T^F})$, but the payment is fixed at zero for output levels below $-\alpha_i^F$ and at $\alpha_i^F + f|T^F(a^F_{T^F})$ for output levels above $f|T^F(a^F_{T^F})$. More formally, for each $i \in T^F$, we define the payment function $\bar{s}_i^F$ as follows:

$$(11)\quad \bar{s}_i^F(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq -\alpha_i^F \\ s_i^F(x) & \text{if } -\alpha_i^F < x < f|T^F(a^F_{T^F}) \\ \alpha_i^F + f|T^F(a^F_{T^F}) & \text{if } x \geq f|T^F(a^F_{T^F}). \end{cases}$$

It is obvious that the payment function $\bar{s}_i^F$ is continuous, piecewise linear, and nondecreasing. Also, it can be easily verified that the NGP $\{T^F, (\bar{s}_i^F)_{i\in T^F}\}$ belongs to $\mathcal{G}_+$.

Clearly, for each $i \in T^F$, the payment according to $\bar{s}_i^F$ can exceed the payment according to s_i^F only at output levels below $-\alpha_i^F$. However, the payment according to $\bar{s}_i^F$ for any output level below $-\alpha_i^F$, which is fixed at zero, is no larger than u_i. Also, for each $i \in T^F$, $\bar{s}_i^F$ and s_i^F award the same payment at the output level $f|T^F(a^F_{T^F})$. Then, as we already know that $(s_i^F)_{i\in T^F}$ induces $a^F_{T^F}$, $(\bar{s}_i^F)_{i\in T^F}$ must also induce $a^F_{T^F}$, and hence, we have the following proposition.

Proposition 3 : *If assumptions* **A1-A5** *hold, and* $B(T^F) \geq V(T^*)$, *then* $(\{T^F, (\bar{s}_i^F)_{i\in T^F}\}, a^F_{T^F}) \in \Omega^*$.

Proof: See Appendix A.

Next, consider the other case, $V(T^*) > B(T^F)$. Partition the set of agents T^*_{-N} into the three subsets, T^*_+, T^*_0 and T^*_-, such that $i \in T^*_+$ if and only if $\alpha_i^* > 0$, $i \in T^*_0$ if and only if $\alpha_i^* = 0$, and $i \in T^*_-$ if and only if $\alpha_i^* < 0$. By relabeling the agents if necessary, without loss of generality, we let T^*_+ be the first $|T^*_+|$ agents, T^*_0 be the $|T^*_0|$ agents after $|T^*_+|$, and T^*_- be the $|T^*_-|$ agents after $|T^*_+| + |T^*_0|$; *i.e.* $T^*_+ = \{1, ..., |T^*_+|\}$ if $T^*_+ \neq \emptyset$, $T^*_0 = \{|T^*_+| + 1, ..., |T^*_+ \cup T^*_0|\}$ if $T^*_0 \neq \emptyset$, and $T^*_- = \{|T^*_+ \cup T^*_0| + 1, ..., |T^*_{-N}|\}$ if $T^*_- \neq \emptyset$.

For each $i \in T^*_-$, let $\bar{s}_i^*$ be the payment function that pays zero wherever s_i^* pays a nonpositive amount and the same as s_i^* everywhere else. Formally, for each $i \in T^*_-$, as the critical output level at which s_i^* starts paying nonnegative amounts is $-\alpha_i^*/\beta_i^*$, we have

$$(12)\quad \bar{s}_i^*(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq -\alpha_i^*/\beta_i^* \\ s_i^*(x) & \text{if } x > -\alpha_i^*/\beta_i^*. \end{cases}$$

For each $i \in T_0^*$, let

(13) $\bar{s}_i^*(x) = s_i^*(x) \quad \forall\, x \in \Re_+$.

Suppose T_+^* is nonempty. Then agent 1 belongs to T_+^*. Now, if agent 1 is paid $x - \sum_{i \in T_0^* \cup T_-^*} \bar{s}_i^*(x)$ for every $x \in \Re_+$, then it is obvious that there is an output level below which she does not get as much as in s_1^* but above which she gets more than in s_1^*. So, let $\bar{x}^1$ be the unique output level such that $s_1^*(\bar{x}^1) = \bar{x}^1 - \sum_{i \in T_0^* \cup T_-^*} \bar{s}_i^*(\bar{x}^1)$. Then the payment function $\bar{s}_1^*$ is defined as follows:

(14) $$\bar{s}_1^*(x) = \begin{cases} x - \sum_{i \in T_0^* \cup T_-^*} \bar{s}_i^*(x) & \text{if } 0 \le x \le \bar{x}^1 \\ s_1^*(x) & \text{if } x > x^1. \end{cases}$$

Following a similar procedure as above, for any $i \in T_+^* - \{1\}$, $\bar{x}^i$ is iteratively defined as the unique critical output level such that $s_i^*(\bar{x}^i) = \bar{x}^i - \sum_{j \in T_0^* \cup T_-^*} \bar{s}_j^*(\bar{x}^i) - \sum_{j=1}^{i-1} \bar{s}_j^*(\bar{x}^i)$. Then, for each $i \in T_+^* - \{1\}$, the payment function $\bar{s}_i^*$ is iteratively defined as follows:

(15) $$\bar{s}_i^*(x) = \begin{cases} x - \sum_{j \in T_0^* \cup T_-^*} \bar{s}_j^*(x) - \sum_{j=1}^{i-1} \bar{s}_j^*(x) & \text{if } 0 \le x \le \bar{x}^i \\ s_i^*(x) & \text{if } x > \bar{x}^i. \end{cases}$$

Note that, if T_+^* has more than one agent, there is an asymmetry in the behaviours of $\bar{s}_1^*$ at output levels on or below $\bar{x}^1$ and $\bar{s}_i^*$ at output levels on or below $\bar{x}^i$, where $i \in T_+^*$ is distinct from 1. For each $i \in T_+^*$ distinct from 1, $\bar{s}_i^*$ pays zero at every output level lower than $\bar{x}^{i-1}$, the critical output level of the agent just before i. On the otherhand, if T_+^* is nonempty, then $\bar{s}_1^*$ pays zero only at zero output level. For each $i \in T_+^*$ distinct from 1, when the final output is on the interval $[\bar{x}^{i-1}, \bar{x}^i]$, $\bar{s}_i^*$ pays agent i the output that is left after paying each agent h before her according to s_h^* and each agent $j \in T_0^* \cup T_-^*$ according to $\bar{s}_j^*$. So, if T_+^* has more than one agent, for each $i \in T_+^*$ distinct from 1, the behaviour of $\bar{s}_i^*$ between $\bar{x}^{i-1}$ and $\bar{x}^i$ is similar to that of $\bar{s}_1^*$ between 0 and $\bar{x}^1$.

It can be easily verified that, for each $i \in T_{-N}^*$, $\bar{s}_i^*$ is continuous, piecewise linear, and nondecreasing. Furthermore, the principal's residual, $x - \sum_{i \in T_{-N}^*} \bar{s}_i^*(x)$, is continuous, piecewise linear, and nondecreasing. Our construction also ensures that the NGP $\{T^*, (\bar{s}_i^*)_{i \in T_{-N}^*}\}$ belongs to $\mathcal{G}_+$.

Clearly, for each $i \in T_+^* \cup T_0^*$, the curve of $\bar{s}_i^*$ always lies on or below that of s_i^*. On the otherhand, for each $i \in T_-^*$, wherever the curve of $\bar{s}_i^*$ lies above that of s_i^* its value is equal to zero, and hence, no larger than the utility of agent i at $a_{T^*}^*$, u_i. Also, the curves of the principal's residual in $(\bar{s}_i^*)_{i \in T_{-N}^*}$ and $(s_i^*)_{i \in T_{-N}^*}$ are such that, if there are output levels at which the former lies above the later, then the value of the former is equal to zero at those output levels. Then, because $(s_i^*)_{i \in T_{-N}^*}$ induces $a_{T^*}^*$, $(\bar{s}_i^*)_{i \in T_{-N}^*}$ should also induce $a_{T^*}^*$.

Proposition 4 : *If assumptions* **A1-A5** *hold, and* $V(T^*) > B(T^F)$, *then* $(\{T^*, (s_i^*)_{i \in T_{-N}^*}\}, a_{T^*}^*) \in \Omega^*$.

Proof. See Appendix A.

Thus, according to Propositions 3 and 4, except for the fact that the principal may have to look for slightly more sophisticated output sharing rules than those of the linear variety (namely, piecewise linear rules), there are no other significant changes when individuals have limited liabilities. This result, as we have pointed out all along, is a consequence of the deterministic production process. In contrast, when the production process is no longer deterministic, often, there is not enough freedom to modify the unlimited liability optimal output sharing rule to a limited liability optimal output sharing rule. Hence, imposing limited liability condition often reduces the principal's optimal utility when there are uncertainties in the production process.

6 Uniqueness of SPE Payoffs

To begin with, we must point out that the result presented in this section relies on the quasilinearity of the utility functions, and hence, may not hold for the more general utility functions that are additively separable and concave but not necessarily quasilinear.

Our objective is to show that the utility tuple remains the same in every SPE outcome of the SBG. Precisely, we show that at any SPE outcome the utility of each agent i is u_i and the utility of the principal is her utility from any OO.

Suppose $(\{T,(s_i)_{i\in T_{-N}}\},a_T)\in\Omega^*$, but the utility of some agent $j\in T_{-N}$, $s_j(f|T(a_T))-c_j(a_j)$, is greater than u_j. Now, if we keep the payment function of every other agent intact and give agent j an $\epsilon>0$ less for every level of output, where ϵ is such that $s_j(f|T(a_T))-\epsilon-c_j(a_j)\geq u_j$, then, because of the quasilinear utility functions, a_T is still a Nash equilibrium at which every agent in T_{-N} (including agent j) get at least their outside option utility and the principal's utility has increased by ϵ. But this means that there is an outcome in $\hat{\Omega}$ at which the principal's utility is higher than at an outcome in Ω^*, which is not possible. So we claim the following.

Claim 3 : *If assumptions* **A1-A5** *hold, and* $(\{T,(s_i)_{i\in T_{-N}}\},a_T)\in\Omega^*$, *then* $s_i(f|T(a_T))-c_i(a_i)=u_i\ \forall\ i\in T_{-N}$.

Suppose $B(T^F)\geq V(T^*)$ ($V(T^*)>B(T^F)$). Then, as assumption **A3** implies $B(T^F)>u_N\geq 0$ ($V(T^*)>u_N\geq 0$), let $\epsilon>0$ be such that $B(T^F)-|T^F|\epsilon>u_N$ ($V(T^*)-|T^*_{-N}|\epsilon>u_N$). Consider the NGP $\{T^F,(s_i^{F\epsilon})_{i\in T^F}\}$ ($\{T^*,(s_i^{*\epsilon})_{i\in T^*_{-N}}\}$), which is obtained from $\{T^F,(s_i^F)_{i\in T^F}\}$ ($\{T^*,(s_i^*)_{i\in T^*_{-N}}\}$) by paying each agent in T^F (T^*_{-N}) ϵ more for every level of output. Then, because of the quasilinear utility functions and Proposition 1 (Proposition 2), it obviously follows that: (i) each agent $i\in T^F$ ($\in T^*_{-N}$) gets $u_i+\epsilon$ at the outcome $(\{T^F,(s_i^{F\epsilon})_{i\in T^F}\},a^F_{T^F})$ $((\{T^*,(s_i^{*\epsilon})_{i\in T^*_{-N}}\},a^*_{T^*}))$; (ii) the principal gets $B(T^F)-|T^F|\epsilon$ ($V(T^*)-|T^*_{-N}|\epsilon$) at the outcome $(\{T^F,(s_i^{F\epsilon})_{i\in T^F}\},a^F_{T^F})$ $((\{T^*,(s_i^{*\epsilon})_{i\in T^*_{-N}}\},a^*_{T^*}))$; and (iii) $a^F_{T^F}$ ($a^*_{T^*}$) is the unique Nash equilibrium of $\{T^F,(s_i^{F\epsilon})_{i\in T^F}\}$ ($\{T^*,(s_i^{*\epsilon})_{i\in T^*_{-N}}\}$).

In the case of limited liabilities, we can exploit the freedom provided by the deterministic production process to modify the above mentioned NGPs in such a way that the limited liability condition is met without loosing any of the conclusions drawn. Thus, Appendix B proves the following lemmas.

Lemma 7 : *Suppose assumptions* **A1-A5** *hold, and* $B(T^F)\geq V(T^*)$. *Then, for each* $\epsilon>0$ *such that* $B(T^F)-|T^F|\epsilon>u_N$, *there exists* $\{T^F,(\hat{s}_j^{F\epsilon})_{j\in T^F}\}\in\mathcal{G}_+$ *such that:* (i) $a^F_{T^F}$ *is its unique Nash equilibrium; and* (ii) *the utilities of the principal and each agent* $i\in T^F$ *at* $a^F_{T^F}$ *are* $B(T^F)-|T^F|\epsilon$ *and* $u_i+\epsilon$, *respectively.*

Proof: See Appendix B.

Lemma 8 : *Suppose assumptions* **A1-A5** *hold, and* $V(T^*)>B(T^F)$. *Then, for each* $\epsilon>0$ *such that* $V(T^*)-|T^*_{-N}|\epsilon>u_N$ *and* $\alpha_i^*+\epsilon<0\ \forall\ i\in T^*_-$ *(the existence of such an* ϵ *is assured by the fact that* $\alpha_i^*<0\ \forall\ i\in T^*_-$*), there exists* $\{T^*,(\hat{s}_j^{*\epsilon})_{j\in T^*_{-N}}\}\in\mathcal{G}_+$ *such that:* (i) $a^*_{T^*}$ *is its unique Nash equilibrium; and* (ii) *the utilities of the principal and each agent* $i\in T^*_{-N}$ *at* $a^*_{T^*}$ *are* $V(T^*)-|T^*_{-N}|\epsilon$ *and* $u_i+\epsilon$, *respectively.*

Proof: See Appendix B.

Now, suppose there is a SPE outcome which does not belong to Ω^*, and the principal's utility at this outcome is $\tilde{u}_N$. Clearly, $\tilde{u}_N$ is less than the principal's utility at an OO. Then, because of Lemmas 7 and 8, we can find some outcome, say $(\{T, (s_i)_{i \in T_{-N}}\}, a_T)$, such that: (i) $\{T, (s_i)_{i \in T_{-N}}\} \in \mathcal{G}_+$; (ii) a_T is the unique Nash equilibrium of $\{T, (s_i)_{i \in T_{-N}}\}$; (iii) T_{-N} is nonempty and the utility of each agent in T_{-N} at a_T is more than her outside option utility; and (iv) the principal's utility at a_T is more than $\tilde{u}_N$. But then, using a simple backward induction logic, it is easy to see that all the agents in T_{-N} will always agree to the NGP $\{T, (s_i)_{i \in T_{-N}}\}$ whenever the principal announces it in the first stage of the SBG. So there cannot be any SPE outcome at which the principal's utility is $\tilde{u}_N$, and hence, our original supposition must be false. Therefore, there are no SPE outcomes outside Ω^*. The uniqueness of the SPE utility tuple is then an immediate consequence of Claims 1 and 3.

Claim 4 : *Suppose assumptions* **A1-A5** *hold. Then, whether there is limited liability or not, we have the following:* (i) *if* $B(T^F) \geq V(T^*)$, *then at any* SPE *the utility of the principal is* $B(T^F)$ *and the utility of each agent* i *is* u_i; *and* (ii) *if* $V(T^*) > B(T^F)$, *then at any* SPE *the utility of the principal is* $V(T^*)$ *and the utility of each agent* i *is* u_i.

7 An Example

Consider a situation with three individuals ($N = 3$). So individuals 1 and 2 are the agents, and individual 3 is the principal. The joint production process and the cost functions are given by: (i) $f(a) = 2(1/6+a_1)^{1/2}(1/6+K_2a_2 + K_3a_3)^{1/2} - 1/3 \; \forall \; a = (a_1, a_2, a_3) \in \Re^3_+$, where K_2 and K_3 are constants to be specified; and (ii) $c_i(a_i) = a_i^2/2 \; \forall \; a_i \in \Re_+$, $i = 1, 2, 3$. All individuals have the same outside option utility, which is equal to 1/3; *i.e.* $u_i = 1/3$, $i = 1, 2, 3$. The constants K_2 and K_3 can be interpreted as parameters that express the relative efficiency between the action of agent 2 and the action of the principal. We look at three different scenarios corresponding to different values of the efficiency parameters K_2 and K_3.

<u>*Case 1*</u>: $K_3 = 1$, and $K_2 > 0$ but sufficiently close to zero.
In this case, it can be easily verified that $B(\{1\}) = B(\{3\}) = 5/24$ and $B(\{1,3\}) = 2/3$. We can also find K_2 small enough such that $B(\{2\}) < 5/24$, $B(\{1,2\}) < 5/24$, $B(\{2,3\}) < 5/24$, and $B(\{1,2,3\}) < 5/12$. Then it is easy to see that the principal must take part in production along with agent 1 to get the maximum utility in the FB situation; *i.e.* $u_3^F = B(\{1,3\}) = 2/3$. So the principal cannot sustain efficiency in this case. Clearly, $T^F = \{1\}$ and $B(T^F) = 5/24$. Also, straightforward maximization shows that $V(\{1,3\}) = 5/12$, and hence, $T^* = \{1,3\}$. Thus, $B(T^F) < V(T^*)$ in this case. Therefore, only agent 1 and the principal participates at an OO. The optimal actions are $(a_1^*, a_3^*) = (1/2, 1/2)$, and the optimal linear and piecewise payment functions for agent 1 are given by:

$$s_1^*(x) = -1/24 + (1/2)x \quad \forall \; x \in \Re_+; \text{ and}$$
$$\bar{s}_1^*(x) = \begin{cases} 0 & \text{if } x \leq 1/12 \\ -1/24 + (1/2)x & \text{if } x > 1/12. \end{cases}$$

<u>*Case 2*</u>: $K_2 = 1$, and $K_3 > 0$ but sufficiently close to zero.
Here, $B(\{1\}) = B(\{2\}) = 5/24$ and $B(\{1,2\}) = 2/3$. We can also find K_3 small enough such that $B(\{3\}) < 5/24$, $B(\{1,3\}) < 5/24$, $B(\{2,3\}) < 5/24$, and $B(\{1,2,3\}) < 2/3$. Then it is obvious that only agents 1 and 2 must take part in production for the principal to get the maximum utility in the FB situation; *i.e.* $u_3^F = B(\{1,2\}) = 2/3$. So $T^F = \{1,2\}$, and $B(T^F) = u_3^F$. Therefore, the principal can sustain efficiency in this case. Clearly, whatever be the T^*, we have $B(T^F) > V(T^*)$. Hence, at an OO, only the two agents participate in production and moral hazard is completely mitigated. It can be easily verified that the optimal actions are

$(a_1^F, a_2^F) = (1,1)$. Thus, the optimal linear and piecewise payment functions for the two agents are given by:

$$s_1^F(x) = s_2^F(x) = -7/6 + x \quad \forall x \in \Re_+; \text{ and}$$
$$\bar{s}_1^F(x) = \bar{s}_2^F(x) = \begin{cases} 0 & \text{if } x \leq 7/6 \\ -7/6 + x & \text{if } 7/6 < x \leq 2 \\ 5/6 & \text{if } x > 2. \end{cases}$$

<u>*Case 3*</u>: $K_2 = 1$, and $K_3 > 1$ but sufficiently close to one.
As in case 2 above, $B(\{1\}) = B(\{2\}) = 5/24$ and $B(\{1,2\}) = 2/3$. Now, we can find K_3 close enough to one such that $B(\{3\}) < 2/3$, $B(\{1,3\}) > 2/3$, $B(\{2,3\}) < 1/3$, $B(\{1,2,3\}) < B(\{1,3\}$, and $V(\{1,2,3\}) < V(\{1,3\})$ $< 2/3$. Then it is obvious that $B(\{1,3\}) \geq B(T)\ \forall\ T \subseteq \{1,2,3\}$. So only agent 1 and the principal must take part in production for the principal to get the maximum utility in the FB situation; *i.e.* $u_3^F = B(\{1,3\}) > 2/3$. It is also easy to see that $T^F = \{1,2\}$, and $T^* = \{1,3\}$. This means that $u_3^F > B(T^F) > V(T^*)$. Therefore, at an OO, as in case 2, only the two agents participate in production and moral hazard is completely mitigated. Moreover, the optimal actions and the linear and piecewise linear optimal payment functions of the two agents remain the same as in case 2. However, unlike case 2, the principal can no longer sustain efficiency as her best option in the FB situation requires her participation in production.

8 Nonquasilinear Utilities

For each agent i, when she participates in production and takes an action a_i and receives a payment m_i, suppose her utility is given by $U_i(m_i, a_i) = v_i(m_i) - c_i(a_i)$, where v_i is concave and satisfies all the other standard assumptions. Similarly, when the principal participates in production and takes an action a_N and receives a residual r, suppose her utility is given by $U_N(r, a_N) = v_N(r) - c_N(a_N)$, where v_N is concave and satisfies all the other standard assumptions. On the otherhand, when the principal does not participate in production but exercises her outside option and receives a residual r, suppose her utility is given by $\tilde{U}_N(r)$, where $\tilde{U}_N$ is concave and satisfies all the other standard assumptions.

With the above specified utilities, the logic about the deterministic production process leaving sufficient room that allows the optimal output sharing rules to behave quite arbitrarily away from the optimal output level is still applicable. Therefore, although it is slightly more demanding technically, we can derive counterparts of conditions (1)-(3) that do not depend on any output sharing rules. Also, for any T such that $N \in T$ and $|T| \geq 2$, we can eliminate the output sharing rule from the appropriate counterpart of problem (P_T) and transform it into the appropriate counterpart of problem (P_T'). Thus, except for the section on the uniqueness of SPE utility tuple (namely, section 6) which relies heavily on the quasilinearity of the utility functions, the analysis in the rest of the paper can be repeated with the more general utility functions without any qualitative changes in the results.

9 Conclusion

We looked at a simple moral hazard problem in a principal-agent(s) framework. However, unlike most existing work, our principal was not precluded from active participation in the production process. Also, unlike the single agent case, there was no uncertainty and the moral hazard problem was caused by joint production. A simple multi-stage extensive game, the SBG, determined the set of individuals who actually took part in production along with the output sharing rule they followed.

Although the principal was not precluded from participation in the production process, whether it was optimal for her to participate or not depended on the values of $B(T^F)$ and $V(T^*)$. In particular, it was best for her to participate only if $V(T^*) \geq B(T^F)$.

Whenever the principal did not participate in production, moral hazard was completely mitigated although there was potential for moral hazard if two or more agents participated in production. On the otherhand, it was impossible to mitigate the moral hazard problem completely if the principal participated in production along with at least one agent. These findings, as we argued, depend on the deterministic production process and not on the quasilinear utility functions. In contrast, unless there is risk neutrality, moral hazard cannot be completely mitigated in most principal-agent moral hazard problems with uncertainty.

From the above remarks we can also draw some other interesting conclusions. Firstly, although the principal could completely mitigate moral hazard by not participating in production, it is quite conceivable that she could be better off introducing moral hazard by participating in production. Secondly, even if it was optimal for the principal to completely mitigate moral hazard by not participating in production, she might still be worse off than at her best option when actions are observable, because her best option when actions are observable might require her participation in the production process.

Except when agents are risk neutral, in most standard principal-agent moral hazard models with nondeterministic production processes, it is the norm rather than the exception that the principal has to look for output sharing rules that are much more sophisticated than linear or piecewise linear output sharing rules. However, although we did not present it formally for the more general additively separable concave utility case, we showed that the principal need not look any further than the class of linear output sharing rules (piecewise linear output sharing rules in case of limited liability) if the production process is deterministic.

Also, in principal-agent moral hazard models with nondeterministic production processes, the results that are obtained without limited liability may change significantly when there is limited liability, for example, the principal's optimal utility often decreases when limited liability is imposed. But we showed that the deterministic production process made most of our results robust to the introduction of limited liability.

Appendix A

Proof of **Lemma 1**: Let $T \in \mathcal{N}$ and $(m_T, a_T) \in \Re^{|T|} \times A^T$ such that $N \notin T$, $m_i - c_i(a_i) \geq u_i \ \forall\ i \in T$ and $u_N + f|T(a_T) - \sum_{i \in T} m_i = B(T)$.

Suppose $m_j - c_j(a_j) > u_j$ for some $j \in T$. Then let $\epsilon > 0$ be such that $m_j - \epsilon - c_j(a_j) \geq u_j$, and define $\hat{m}_T \in \Re^{|T|}$ as $\hat{m}_i = m_i \ \forall\ i \in T - \{j\}$ and $\hat{m}_j = m_j - \epsilon$. Clearly, $(\hat{m}_T, a_T)$ is such that $\hat{m}_i - c_i(a_i) \geq u_i \ \forall\ i \in T$ and $u_N + f|T(a_T) - \sum_{i \in T} \hat{m}_i > B(T)$, a contradiction to the definition of $B(T)$. Hence, (i) of Lemma 1 holds.

(i) of Lemma 1 implies that $u_N + f|T(a_T) - \sum_{i \in T} c_i(a_i) - \sum_{i \in T} u_i = B(T)$. Now, suppose there exists $\tilde{a}_T \in A^T$ such that $u_N + f|T(\tilde{a}_T) - \sum_{i \in T} c_i(\tilde{a}_i) - \sum_{i \in T} u_i > u_N + f|T(a_T) - \sum_{i \in T} c_i(a_i) - \sum_{i \in T} u_i$. Let $\tilde{m}_T \in \Re^{|T|}$ be such that $\tilde{m}_i = u_i + c_i(\tilde{a}_i) \ \forall\ i \in T$. Then it is obvious that $(\tilde{m}_T, \tilde{a}_T)$ satisfies $\tilde{m}_i - c_i(\tilde{a}_i) = u_i \ \forall\ i \in T$ and $u_N + f|T(\tilde{a}_T) - \sum_{i \in T} \tilde{m}_i > B(T)$, which contradicts the definition of $B(T)$. Thus, a_T solves the problem

$$\max_{a'_T \in A^T} [u_N + f|T(a'_T) - \textstyle\sum_{i \in T} c_i(a'_i) - \sum_{i \in T} u_i].$$

Because of the limiting properties of the derivatives of c_i and f given in assumptions **A1** and **A2**, the above maximization problem can only have interior solutions. So (ii) of Lemma 1 must hold. Also, as an interior solution of the above maximization problem, a_T must satisfy the first order conditions, $f_i|T(a_T) - c'_i(a_i) = 0 \ \forall$ $i \in T$, which are exactly the conditions in (iii) of Lemma 1. ||

Proof of **Lemma 2**: Similar to the proof of Lemma 1. ||

Proof of **Proposition 1**: (ii) and (iii) of Proposition 1 readily follow from (4) and (5). Also, it is obvious from (5) that $s_i^F \in S \ \forall\ i \in T^F$. So, if we show (iv) of Proposition 1, then we have also shown (i) of Proposition 1.

The strict convexity of c_i, the concavity of f and (5) imply that $s_i^F(f|T^F(a_{T^F})) - c_i(a_i)$ is concave in $a_{T^F} \in A^{T^F}$ for each $i \in T^F$. Thus, because of (5), if $\tilde{a}_{T^F} \in A^{T^F}$ satisfies $f_i|T^F(\tilde{a}_{T^F}) - c'_i(\tilde{a}_i) = 0 \ \forall\ i \in T^F$, then $\tilde{a}_{T^F} \in NE(\{T^F, (s_i^F)_{i \in T^F}\})$. But we already know from Lemma 1 that $f_i|T^F(a^F_{T^F}) - c'_i(a_i^F) = 0 \ \forall\ i \in T^F$. So $a^F_{T^F} \in NE(\{T^F, (s_i^F)_{i \in T^F}\})$.

Using (5) and the limiting properties of the derivatives of c_i and f given in assumptions **A1** and **A2**, it is quite obvious that, at any Nash equilibrium of $\{T^F, (s_i^F)_{i \in T^F}\}$, the actions of all the agents in T^F are positive. Thus, in fact any $\tilde{a}_{T^F} \in A^{T^F}$ is a Nash equilibrium of $\{T^F, (s_i^F)_{i \in T^F}\}$ if and only if $f_i|T^F(\tilde{a}_{T^F}) - c'_i(\tilde{a}_i) = 0 \ \forall$ $i \in T^F$.

Now, consider the following maximization problem:

$$\text{(A1)} \qquad \max_{a_{T^F} \in A^{T^F}} [f|T^F(a_{T^F}) - \textstyle\sum_{i \in T^F} c_i(a_i)].$$

The concavity of f and the strict convexity of c_i imply that the objective function of the maximization problem in (A1) is strictly concave in $a_{T^F} \in A^{T^F}$. Also, because of the limiting properties of the derivatives of c_i and f given in assumptions **A1** and **A2**, the problem in (A1) can only have interior solutions. Thus, $\tilde{a}_{T^F} \in A^{T^F}$ is a solution of the problem in (A1) if and only if it satisfies $f_i|T^F(\tilde{a}_{T^F}) - c'_i(\tilde{a}_i) = 0 \ \forall\ i \in T^F$, which are the first order conditions. But this immediately implies that $\tilde{a}_{T^F} \in A^{T^F}$ is a Nash equilibrium of $\{T^F, (s_i^F)_{i \in T^F}\}$ if and only if it is also a solution of the problem in (A1). However, the problem in (A1) can have at the most one solution, because we already know that its objective function is strictly convave. Therefore $\{T^F, (s_i^F)_{i \in T^F}\}$ can have at the most one Nash equilibrium. Hence, because $a^F_{T^F} \in NE(\{T^F, (s_i^F)_{i \in T^F}\})$, (iv) of Proposition 1 must hold as well. ||

Proof of **Lemma 3**: Let $T \in \mathcal{N}$ be such that $N \in T$ and $|T| \geq 2$. Suppose $(\{T, (s_j)_{j \in T_{-N}}\}, a_T)$ satisfies (C1) and (C2). Let i be any member of T_{-N}. Then the proof of Lemma 3 is completed in two steps. In the first step we show that s_i is continuous at $f|T(a_T)$. The continuity of s_j at $f|T(a_T)$ for every $j \in T_{-N}$ is then used in the second step to show that s_i is differentiable at $f|T(a_T)$.

Step 1:

To prove that s_i is continuous at $f|T(a_T)$ (> 0) it is sufficient to show that

$$\text{(A2)}\quad \lim_{z\uparrow 0} s_i(f|T(a_T)+z) \;=\; \lim_{z\downarrow 0} s_i(f|T(a_T)+z) \;=\; s_i(f|T(a_T)),$$

where $z \uparrow 0$ and $z \downarrow 0$ denote z approaching 0 through negative and positive values, respectively.

Clearly, constraints (C1) and (C2) imply that, if $y \in \Re$ and $a_i + y \in A_i$, then $s_i(f|T(a_{T_{-i}}, a_i+y)) - c_i(a_i+y) \leq s_i(f|T(a_T)) - c_i(a_i)$, and hence, neither of $\lim_{y\uparrow 0}[s_i(f|T(a_{T_{-i}}, a_i+y)) - c_i(a_i+y)]$ or $\lim_{y\downarrow 0}[s_i(f|T(a_{T_{-i}}, a_i+y)) - c_i(a_i+y)]$ can be greater than $s_i(f|T(a_T)) - c_i(a_i)$. Then, because of the continuity of c_i, we have

$$\lim_{y\uparrow 0} s_i(f|T(a_{T_{-i}}, a_i+y)) \leq s_i(f|T(a_T)), \quad \text{and} \quad \lim_{y\downarrow 0} s_i(f|T(a_{T_{-i}}, a_i+y)) \leq s_i(f|T(a_T)).$$

However, because of the continuity and monotonicity of f, we also have

$$\lim_{y\uparrow 0} s_i(f|T(a_{T_{-i}}, a_i+y)) \;=\; \lim_{z\uparrow 0} s_i(f|T(a_T)+z), \quad \text{and}$$
$$\lim_{y\downarrow 0} s_i(f|T(a_{T_{-i}}, a_i+y)) \;=\; \lim_{z\downarrow 0} s_i(f|T(a_T)+z).$$

Thus, the following must be true:

$$\text{(A3)}\quad \lim_{z\uparrow 0} s_i(f|T(a_T)+z) \;\leq\; s_i(f|T(a_T)), \quad \text{and}$$
$$\text{(A4)}\quad \lim_{z\downarrow 0} s_i(f|T(a_T)+z) \;\leq\; s_i(f|T(a_T)).$$

(C1) and (C2) also imply that, if $y \in \Re$ and $a_N + y \in A_N$, then $f|T(a_{T_{-N}}, a_N+y) - \sum_{j\in T_{-N}} s_j(f|T(a_{T_{-N}}, a_N+y)) - c_N(a_N+y) \leq f|T(a_T) - \sum_{j\in T_{-N}} s_j(f|T(a_T)) - c_N(a_N)$, and hence, neither of $\lim_{y\uparrow 0}[f|T(a_{T_{-N}}, a_N+y) - \sum_{j\in T_{-N}} s_j(f|T(a_{T_{-N}}, a_N+y)) - c_N(a_N+y)]$ or $\lim_{y\downarrow 0}[f|T(a_{T_{-N}}, a_N+y) - \sum_{j\in T_{-N}} s_j(f|T(a_{T_{-N}}, a_N+y)) - c_N(a_N+y)]$ can be greater than $f|T(a_T) - \sum_{j\in T_{-N}} s_j(f|T(a_T)) - c_N(a_N)$. Then, because of the continuity of c_N and f, we have

$$\textstyle\sum_{j\in T_{-N}} \lim_{y\uparrow 0} s_j(f|T(a_{T_{-N}}, a_N+y)) \;\geq\; \sum_{j\in T_{-N}} s_j(f|T(a_T)), \quad \text{and}$$
$$\textstyle\sum_{j\in T_{-N}} \lim_{y\downarrow 0} s_j(f|T(a_{T_{-N}}, a_N+y)) \;\geq\; \sum_{j\in T_{-N}} s_j(f|T(a_T)).$$

However, for each $j \in T_{-N}$, because of the continuity and monotonicity of f, we also have

$$\lim_{y\uparrow 0} s_j(f|T(a_{T_{-N}}, a_N+y)) \;=\; \lim_{z\uparrow 0} s_j(f|T(a_T)+z), \quad \text{and}$$
$$\lim_{y\downarrow 0} s_j(f|T(a_{T_{-N}}, a_N+y)) \;=\; \lim_{z\downarrow 0} s_j(f|T(a_T)+z).$$

Hence, the following must hold:

$$\text{(A5)}\quad \textstyle\sum_{j\in T_{-N}} \lim_{z\uparrow 0} s_j(f|T(a_T)+z) \;\geq\; \sum_{j\in T_{-N}} s_j(f|T(a_T)), \quad \text{and}$$
$$\text{(A6)}\quad \textstyle\sum_{j\in T_{-N}} \lim_{z\downarrow 0} s_j(f|T(a_T)+z) \;\geq\; \sum_{j\in T_{-N}} s_j(f|T(a_T)).$$

Now, (A2) readily follows from (A3)-(A6). Therefore, s_i is continuous at $f|T(a_T)$.

Step 2:

Let $x^o = f|T(a_T)$. Because of (C1) and (C2), s_i is piecewise continuous and $x^o > 0$. Then, as s_i is continuous at x^o, there exists $\hat{\epsilon} > 0$ sufficiently small such that $x^o - \hat{\epsilon} > 0$ and s_i is continuous on $(x^o - \hat{\epsilon}, x^o + \hat{\epsilon})$. So, by (ii) in the definition of S, we can find $\hat{\delta} > 0$ such that $\hat{\delta} \leq \hat{\epsilon}$ and s_i is continuously differentiable on the two intervals $(x^o - \hat{\delta}, x^o)$ and $(x^o, x^o + \hat{\delta})$. Let s_i' denote the derivative of s_i wherever it exists. Then to prove that s_i is differentiable at x^o it is sufficient to show that

$$\text{(A7)}\quad \lim_{z\uparrow 0} s_i'(x^o+z) \;=\; \lim_{z\downarrow 0} s_i'(x^o+z).$$

As f and c_i are continuously differentiable, (C1), (C2) and the continuous differentiability of s_i on $(x^o-\hat{\delta}, x^o)$ and $(x^o, x^o+\hat{\delta})$ immediately imply that

$$\text{(A8)} \qquad \lim_{y\uparrow 0} s_i'(f|T(a_{T_{-i}}, a_i+y))f_i|T(a_T) - c_i'(a_i) \geq 0,$$

$$\text{(A9)} \qquad \lim_{y\downarrow 0} s_i'(f|T(a_{T_{-i}}, a_i+y))f_i|T(a_T) - c_i'(a_i) \leq 0,$$

$$\text{(A10)} \quad \lim_{y\uparrow 0}(1-\textstyle\sum_{j\in T_{-N}} s_j'(f|T(a_{T_{-N}}, a_N+y)))f_N|T(a_T) - c_N'(a_N) \geq 0, \text{ and}$$

$$\text{(A11)} \quad \lim_{y\downarrow 0}(1-\textstyle\sum_{j\in T_{-N}} s_j'(f|T(a_{T_{-N}}, a_N+y)))f_N|T(a_T) - c_N'(a_N) \leq 0.$$

Given (C1), for each $j \in T_{-N}$, the continuity and monotonicity of f also imply that

$$\text{(A12)} \quad \lim_{y\uparrow 0} s_j'(f|T(a_{T_{-j}}, a_j+y)) = \lim_{y\uparrow 0} s_j'(f|T(a_{T_{-N}}, a_N+y)) = \lim_{z\uparrow 0} s_j'(x^o+z), \quad \text{and}$$

$$\text{(A13)} \quad \lim_{y\downarrow 0} s_j'(f|T(a_{T_{-j}}, a_j+y)) = \lim_{y\downarrow 0} s_j'(f|T(a_{T_{-N}}, a_N+y)) = \lim_{z\downarrow 0} s_j'(x^o+z).$$

So, by using (A12) and (A13) in (A8)-(A11), we get

$$\text{(A14)} \qquad \lim_{z\uparrow 0} s_i'(x^o+z) \geq \frac{c_i'(a_i)}{f_i|T(a_T)} \geq \lim_{z\downarrow 0} s_i'(x^o+z), \qquad \text{and}$$

$$\text{(A15)} \quad \textstyle\sum_{j\in T_{-N}}[\lim_{z\uparrow 0} s_j'(x^o+z)] \leq 1-\dfrac{c_N'(a_N)}{f_N|T(a_T)} \leq \textstyle\sum_{j\in T_{-N}}[\lim_{z\downarrow 0} s_j'(x^o+z)].$$

Now, (A7) easily follows from (A14) and (A15). Hence, s_i is differentiable at x^o. ||

Proof of **Lemma 4**: Let $T \in \mathcal{N}$ be such that $N \in T$ and $|T| \geq 2$. Suppose $(\{T, (s_j)_{j\in T_{-N}}\}, a_T)$ satisfies (C1) and (C3). Then, because of Lemma 3, (C1) and (C2) immediately imply that

$$\text{(A16)} \qquad s_i'(f|T(a_T))f_i|T(a_T) - c_i'(a_i) = 0 \quad \forall\, i \in T_{-N}, \quad \text{and}$$

$$\text{(A17)} \quad (1-\textstyle\sum_{j\in T_{-N}} s_j'(f|T(a_T)))f_N|T(a_T) - c_N'(a_N) = 0.$$

Also, because of the monotonicity of the cost functions and the production function and (C1), it is obvious that, for each $i \in T$, $c_i'(a_i) > 0$ and $f_i|T(a_T) > 0$. Hence, by some simple algebraic manipulations of (A16) and (A17), we get $\sum_{i\in T}[c_i'(a_i)/f_i|T(a_T)] = 1$. ||

Proof of **Lemma 5**: Let $T \in \mathcal{N}$ be such that $N \in T$ and $|T| \geq 2$. The limiting properties of the derivatives of the cost functions and the production function in assumptions **A1** and **A2** imply $\lim_{\|a_T\|\to 0}\sum_{i\in T}[c_i'(a_i)/f_i|T(a_T)] = 0$ and $\lim_{\|a_T\|\to\infty}\sum_{i\in T}[c_i'(a_i)/f_i|T(a_T)] = \infty$, where $\|.\|$ is the standard Eucleadian norm. Then, because of the continuous differentiability and the monotonicity and curvature properties of the cost functions and the production function in assumptions **A1** and **A2**, it is easily verified that the feasible set of problem (P_T') is nonempty and compact. As the cost functions and the production function are continuously differentiable, it is obvious that the objective function of problem (P_T') is continuous. Hence, the theorem of Weierstrass implies the existence of a solution to problem (P_T'). ||

Proof of **Lemma 6**: Let $T \in \mathcal{N}$ be such that $N \in T$ and $|T| \geq 2$. Suppose $(\{T, (s_j)_{j\in T_{-N}}\}, a_T)$ satisfies (C1) and (C2). Then Lemma 4 implies that a_T is feasible for problem (P_T'). Obviously, because of Lemma 5, $V(T)$ is well defined. Hence, $V(T) \geq f|T(a_T) - \sum_{i\in T} c_i(a_i) - \sum_{j\in T_{-N}} u_j$. But (C2) implies $s_j(f|T(a_T)) - c_j(a_j) \geq u_j$ $\forall$ $j \in T_{-N}$. Therefore, $f|T(a_T) - \sum_{i\in T} c_i(a_i) - \sum_{j\in T_{-N}} u_j \geq f|T(a_T) - \sum_{j\in T_{-N}} s_j(f|T(a_T)) - c_N(a_N)$. Thus, it must be the case that $V(T) \geq f|T(a_T) - \sum_{j\in T_{-N}} s_j(f|T(a_T)) - c_N(a_N)$. ||

Proof of **Proposition 2**: Suppose assumptions **A1-A5** are satisfied and $V(T^*) > B(T^F)$.

(7), (9) and (10) immediately imply (ii) and (iii) of Proposition 2. It is obvious from (10) that $s_i^* \in S$ $\forall$ $i \in T_{-N}^*$. So, if we show that $a_i^* > 0$ $\forall$ $i \in T^*$ and (iv) of Proposition 2 holds, then, because of Lemma 6, the proof of (i) of Proposition 2 is also complete.

Clearly, (6), (8) and the monotonicity of the cost functions and the production function imply that $\beta_i^* \geq 0$ $\forall\, i \in T^*_{-N}$ and $(1 - \sum_{j\in T^*_{-N}} \beta_j^*) \geq 0$. Then, because of the strict convexity of the cost functions, the concavity of the production function and (10), $s_i^*(f|T^*(a_{T^*})) - c_i(a_i)$ is concave in $a_{T^*} \in A^{T^*}$ for each $i \in T^*_{-N}$ and $f|T^*(a_{T^*}) - \sum_{j\in T^*_{-N}} s_j^*(f|T^*(a_{T^*})) - c_N(a_N)$ is concave in $a_{T^*} \in A^{T^*}$. Thus, because of (10), if $\tilde{a}_{T^*} \in A^{T^*}$ satisfies $\beta_i^* f_i|T^*(\tilde{a}_{T^*}) - c_i'(\tilde{a}_i) = 0\ \forall\, i \in T^*_{-N}$ and $(1 - \sum_{j\in T^*_{-N}} \beta_j^*) f_N|T^*(\tilde{a}_{T^*}) - c_N'(\tilde{a}_N) = 0$, then $\tilde{a}_{T^*} \in NE(\{T^*, (s_j^*)_{j\in T^*_{-N}}\})$. Hence, (6) and (8) imply $a^*_{T^*} \in NE(\{T^*, (s_j^*)_{j\in T^*_{-N}}\})$.

Thus, it immediately follows that $(\{T^*, (s_j^*)_{j\in T^*_{-N}}\}, a^*_{T^*}) \in \hat{\Omega}$. But suppose $a_k^* = 0$ for some $k \in T^*$. Then assumptions **A4** and **A5**, and $V(T^*) > B(T^F)$ imply that there is some $(\{T, (s_j)_{j\in T_{-N}}\}, a_T) \in \hat{\Omega}_{+N}$ such that $\pi((\{T, (s_j)_{j\in T_{-N}}\}, a_T)) > V(T^*)$, $a_j > 0\ \forall\, j \in T$, and $|T| \geq 2$. But then, because of Lemma 6, we get $\pi((\{T, (s_j)_{j\in T_{-N}}\}, a_T)) > V(T^*) \geq V(T) \geq \pi((\{T, (s_j)_{j\in T_{-N}}\}, a_T))$, which is impossible. Therefore, $a_i^* > 0$ $\forall\, i \in T^*$.

Having shown that $a_i^* > 0\ \forall\, i \in T^*$, we can immediately conclude that

$$\text{(A18)}\qquad 1 > \beta_i^* > 0 \quad \forall\, i \in T^*_{-N}, \quad \text{and} \quad 1 > (1 - \textstyle\sum_{j\in T^*_{-N}} \beta_j^*) > 0.$$

So, using (10), (A18) and the limiting properties of the derivatives of the cost functions and the production function given in assumptions **A1** and **A2**, it is quite obvious that at any Nash equilibrium of $\{T^*, (s_j^*)_{j\in T^*_{-N}}\}$ the actions of all the individuals in T^* are positive. Thus, in fact any $\tilde{a}_{T^*} \in A^{T^*}$ is a Nash equilibrium of $\{T^*, (s_j^*)_{j\in T^*_{-N}}\}$ if and only if $\beta_i^* = c_i'(\tilde{a}_i)/f_i|T^*(\tilde{a}_{T^*})\ \forall\, i \in T^*_{-N}$ and $(1 - \sum_{j\in T^*_{-N}} \beta_j^*) = c_N'(\tilde{a}_N)/f_N|T^*(\tilde{a}_{T^*})$.

Consider the following maximization problem:

$$\text{(A19)}\qquad \max_{a_{T^*}\in A^{T^*}} [f|T^*(a_{T^*}) - \textstyle\sum_{j\in T^*_{-N}} [c_j(a_j)/\beta_j^*] - (c_N(a_N)/(1 - \sum_{j\in T^*_{-N}} \beta_j^*))].$$

The concavity of the production function, the strict convexity of the cost functions and (A18) imply that the objective function of the maximization problem in (A19) is strictly concave in $a_{T^*} \in A^{T^*}$. Also, because of (A18) and the limiting properties of the derivatives of the cost functions and the production function given in assumptions **A1** and **A2**, the problem in (A19) can only have interior solutions. Thus, $\tilde{a}_{T^*} \in A^{T^*}$ is a solution of the problem in (A19) if and only if it satisfies $\beta_i^* = c_i'(\tilde{a}_i)/f_i|T^*(\tilde{a}_{T^*})\ \forall\, i \in T^*_{-N}$ and $(1 - \sum_{j\in T^*_{-N}} \beta_j^*) = c_N'(\tilde{a}_N)/f_N|T^*(\tilde{a}_{T^*})$, which are the first order conditions. But this immediately implies that $\tilde{a}_{T^*} \in A^{T^*}$ is a Nash equilibrium of $\{T^*, (s_j^*)_{j\in T^*_{-N}}\}$ if and only if it is also a solution of the problem in (A19). However, the problem in (A19) can have at the most one solution, because we already know that its objective function is strictly convave. Therefore, $\{T^*, (s_j^*)_{j\in T^*_{-N}}\}$ can have at the most one Nash equilibrium. Hence, as $a^*_{T^*} \in NE(\{T^*, (s_j^*)_{j\in T^*_{-N}}\})$, (iv) of Proposition 2 must hold. ||

Proof of **Proposition 3**: Using (11), and (ii) and (iii) of Proposition 1, it is quite obvious that

$$\text{(A20)}\qquad \bar{s}_i^F(f|T^F(a^F_{T^F})) - c_i(a_i^F) = u_i \quad \forall\, i \in T^F, \text{ and}$$

$$\text{(A21)}\qquad u_N + f|T^F(a^F_{T^F}) - \sum_{i\in T^F} \bar{s}_i^F(f|T^F(a^F_{T^F})) = B(T^F).$$

Then, given $B(T^F) \geq V(T^*)$, it is sufficient to show that $a^F_{T^F} \in NE(\{T^F, (\bar{s}_i^F)_{i\in T^F}\})$.

For each $i \in T^F$, as c_i is strictly increasing and $c_i(0) = 0$, (11), (A20) and $u_i \geq 0$ imply $\bar{s}_i^F(f|T^F(a^F_{T^F_{-i}}, a_i)) - c_i(a_i) \leq \bar{s}_i^F(f|T^F(a^F_{T^F})) - c_i(a_i^F)$ for any $a_i \in A_i$ such that $0 \leq f|T^F(a^F_{T^F_{-i}}, a_i) \leq -\alpha_i^F$.

For each $i \in T^F$, using (ii) and (iv) of Proposition 1, (11) and (A20), it can be verified that $\bar{s}_i^F(f|T^F(a^F_{T^F_{-i}}, a_i)) - c_i(a_i) \leq \bar{s}_i^F(f|T^F(a^F_{T^F})) - c_i(a_i^F)$ for any $a_i \in A_i$ such that $-\alpha_i^F < f|T^F(a^F_{T^F_{-i}}, a_i) < f|T^F(a^F_{T^F})$.

For each $i \in T^F$, as both c_i and f are strictly increasing in $a_i \in A_i$, it is obvious that, for any $a_i \in A_i$, $f|T^F(a^F_{T^F_{-i}}, a_i) \geq f|T^F(a^F_{T^F})$ if and only if $c_i(a_i) \geq c_i(a_i^F)$. Hence, for each $i \in T^F$, (11) implies that $\bar{s}_i^F(f|T^F(a^F_{T^F_{-i}}, a_i)) - c_i(a_i) \leq \bar{s}_i^F(f|T^F(a^F_{T^F})) - c_i(a_i^F)$ for any $a_i \in A_i$ such that $f|T^F(a^F_{T^F_{-i}}, a_i) \geq f|T^F(a^F_{T^F})$.

Thus, for each $i \in T^F$, we have established that $\bar{s}_i^F(f|T^F(a^F_{T^F_{-i}}, a_i)) - c_i(a_i) \leq \bar{s}_i^F(f|T^F(a^F_{T^F})) - c_i(a_i^F)$ for any $a_i \in A_i$. Therefore, $a^F_{T^F} \in NE(\{T^F, (\bar{s}_i^F)_{i \in T^F}\})$. ||

Proof of **Proposition 4**: Suppose assumptions **A1-A5** are satisfied and $V(T^*) > B(T^F)$.

Because of (ii) and (iii) of Proposition 2, (12)-(15) readily imply that

$$\text{(A22)} \qquad \bar{s}_i^*(f|T^*(a^*_{T^*})) - c_i(a_i^*) \;=\; u_i \quad \forall\, i \in T^*_{-N}, \text{ and}$$

$$\text{(A23)} \qquad f|T^*(a^*_{T^*}) - \sum_{i \in T^*_{-N}} \bar{s}_i^*(f|T^*(a^*_{T^*})) - c_N(a^*_N) \;=\; V(T^*).$$

So it is sufficient to show that $a^*_{T^*} \in NE(\{T^*, (\bar{s}_i^*)_{i \in T^*_{-N}}\})$.

It is obvious from (13) and (iv) of Proposition 2 that, for each $i \in T^*_0$, $\bar{s}_i^*(f|T^*(a^*_{T^*_{-i}}, a_i)) - c_i(a_i) \leq \bar{s}_i^*(f|T^*(a^*_{T^*})) - c_i(a_i^*)\ \forall\, a_i \in A_i$.

For each $i \in T^*_-$, as c_i is strictly increasing and $c_i(0) = 0$, (12), (A22) and $u_i \geq 0$ imply $\bar{s}_i^*(f|T^*(a^*_{T^*_{-i}}, a_i)) - c_i(a_i) \leq \bar{s}_i^*(f|T^*(a^*_{T^*})) - c_i(a_i^*)$ for any $a_i \in A_i$ such that $0 \leq f|T^*(a^*_{T^*_{-i}}, a_i) \leq -\alpha_i^*/\beta_i^*$.

For each $i \in T^*_-$, using (iv) of Proposition 2 and (12), it is easily verified that $\bar{s}_i^*(f|T^*(a^*_{T^*_{-i}}, a_i)) - c_i(a_i) \leq \bar{s}_i^*(f|T^*(a^*_{T^*})) - c_i(a_i^*)$ for any $a_i \in A_i$ such that $f|T^*(a^*_{T^*_{-i}}, a_i) > -\alpha_i^*/\beta_i^*$.

It can be checked in (14) and (15) that, for each $i \in T^*_+$, $\bar{s}_i^*(x) \leq s_i^*(x)\ \forall\, x \in \Re_+$. So, for each $i \in T^*_+$, (ii) and (iv) of Proposition 2 and (A22) imply that $\bar{s}_i^*(f|T^*(a^*_{T^*_{-i}}, a_i)) - c_i(a_i) \leq \bar{s}_i^*(f|T^*(a^*_{T^*})) - c_i(a_i^*)\ \forall\, a_i \in A_i$.

Clearly $V(T^*) > u_N \geq 0$. Now, suppose $|T^*_+| > 0$. Then it is easily verified from (12)-(15) that $x - \sum_{j \in T^*_{-N}} \bar{s}_j^*(x) = 0\ \forall\, x \in [0, \bar{x}^{|T^*_+|}]$ and $x - \sum_{j \in T^*_{-N}} \bar{s}_j^*(x) \leq x - \sum_{j \in T^*_{-N}} s_j^*(x)\ \forall\, x > \bar{x}^{|T^*_+|}$. Then, as c_N is strictly increasing and $c_N(0) = 0$, (A23) and $V(T^*) > 0$ imply that $f|T^*(a^*_{T^*_{-N}}, a_N) - \sum_{j \in T^*_{-N}} \bar{s}_j^*(f|T^*(a^*_{T^*_{-N}}, a_N)) - c_N(a_N) \leq f|T^*(a^*_{T^*}) - \sum_{j \in T^*_{-N}} \bar{s}_j^*(f|T^*(a^*_{T^*})) - c_N(a^*_N)\ \forall\, a_N \in A_N$ such that $0 \leq f|T^*(a^*_{T^*_{-N}}, a_N) \leq \bar{x}^{|T^*_+|}$. Also, (iii) and (iv) of Proposition 2 and (A23) imply that $f|T^*(a^*_{T^*_{-N}}, a_N) - \sum_{j \in T^*_{-N}} \bar{s}_j^*(f|T^*(a^*_{T^*_{-N}}, a_N)) - c_N(a_N) \leq f|T^*(a^*_{T^*}) - \sum_{j \in T^*_{-N}} \bar{s}_j^*(f|T^*(a^*_{T^*})) - c_N(a^*_N)$ for any $a_N \in A_N$ such that $f|T^*(a^*_{T^*_{-N}}, a_N) > \bar{x}^{|T^*_+|}$.

Next, suppose $|T^*_+| = 0$. Then it is clear from (12) and (13) that $x - \sum_{j \in T^*_{-N}} \bar{s}_j^*(x) \leq x - \sum_{j \in T^*_{-N}} s_j^*(x)\ \forall\, x \in \Re_+$. So (iii) and (iv) of Proposition 2 and (A23) imply that $f|T^*(a^*_{T^*_{-N}}, a_N) - \sum_{j \in T^*_{-N}} \bar{s}_j^*(f|T^*(a^*_{T^*_{-N}}, a_N)) - c_N(a_N) \leq f|T^*(a^*_{T^*}) - \sum_{j \in T^*_{-N}} s_j^*(f|T^*(a^*_{T^*})) - c_N(a^*_N)\ \forall\, a_N \in A_N$.

Thus, we have established that, for each $i \in T^*_{-N}$, $\bar{s}_i^*(f|T^*(a^*_{T^*_{-i}}, a_i)) - c_i(a_i) \leq \bar{s}_i^*(f|T^*(a^*_{T^*})) - c_i(a_i^*)$ for any $a_i \in A_i$, and $f|T^*(a^*_{T^*_{-N}}, a_N) - \sum_{j \in T^*_{-N}} \bar{s}_j^*(f|T^*(a^*_{T^*_{-N}}, a_N)) - c_N(a_N) \leq f|T^*(a^*_{T^*}) - \sum_{j \in T^*_{-N}} \bar{s}_j^*(f|T^*(a^*_{T^*})) - c_N(a^*_N)$ for any $a_N \in A_N$. Hence, $a^*_{T^*} \in NE(\{T^*, (\bar{s}_i^*)_{i \in T^*_{-N}}\})$. ||

Appendix B

We first prove Lemma 9 below, which is used in the construction of the output sharing rule used to prove Lemma 7.

Lemma 9 : *Suppose assumptions* **A1** *and* **A2** *are satisfied, and* $\{T^F, (s_i^e)_{i \in T^F}\}$ *is such that* $s_i^e(x) = x/|T^F|\ \forall\, x \in \Re_+$ *and each* $i \in T^F$. *Then there exists* $y^F > 0$ *such that* $y^F < f|T^F(a_{T^F})\ \forall\, a_{T^F} \in NE(\{T^F, (s_i^e)_{i \in T^F}\})$.

Proof of **Lemma 9**: Suppose $\{T^F, (s_i^\epsilon)_{i\in T^F}\}$ is as given in the statement of the lemma. Then the limiting properties of the derivatives of c_i and f given in assumptions **A1** and **A2** imply that, if $a_{T^F} \in A^{T^F}$ is a Nash equilibrium of $\{T^F, (s_i^\epsilon)_{i\in T^F}\}$, then $a_i > 0 \ \forall\ i \in T^F$, and hence, $f|T^F(a_{T^F}) > 0$. So Lemma 9 must hold. ||

Suppose $B(T^F) \geq V(T^*)$. Then it is obvious that $B(T^F) > u_N \geq 0$. Now, pick any $\epsilon > 0$ such that $B(T^F) - |T^F|\epsilon > u_N$. Define $s_i^{F\epsilon}(x) = s_i^F(x) + \epsilon \ \forall\ x \in \Re_+$ and each $i \in T^F$. Using Proposition 1 and the quasilinearity of the utility functions, it can be easily checked that

$$\text{(B1)} \qquad s_i^{F\epsilon}(f|T^F(a_{T^F}^F)) - c_i(a_i^F) \;=\; u_i + \epsilon \quad \forall\ i \in T^F,$$

$$\text{(B2)} \qquad u_N + f|T^F(a_{T^F}^F) - \sum_{i\in T^F} s_i^{F\epsilon}(f|T^F(a_{T^F}^F)) \;=\; B(T^F) - |T^F|\epsilon, \quad \text{and}$$

$$\text{(B3)} \qquad NE(\{T^F, (s_i^{F\epsilon})_{i\in T^F}\}) \;=\; \{a_{T^F}^F\}.$$

Next, let $\underline{y}^\epsilon > 0$ be such that $\underline{y}^\epsilon < \min_{i\in T^F}\{u_i\} + \epsilon$, $\underline{y}^\epsilon < f|T^F(a_{T^F}^F)$, and $\underline{y}^\epsilon < y^F$, where y^F is as given in Lemma 9. Also, let $\underline{y}^\epsilon < \bar{y}^\epsilon < f|T^F(a_{T^F}^F)$. Furthermore, let $\bar{\bar{y}}^\epsilon > f|T^F(a_{T^F}^F)$ be such that $\bar{\bar{y}}^\epsilon - \sum_{i\in T^F} s_i^{F\epsilon}(\bar{\bar{y}}^\epsilon) > 0$. The existence of $\bar{\bar{y}}^\epsilon$ is gauranteed by $B(T^F) - |T^F|\epsilon > u_N$ and (B2). Then, for each $i \in T^F$, define the payment function $\hat{s}_i^{F\epsilon}$ as follows:

$$\text{(B4)} \qquad \hat{s}_i^{F\epsilon}(x) \;=\; \begin{cases} x/|T^F| & \text{if } 0 \leq x < \underline{y}^\epsilon \\ 0 & \text{if } \underline{y}^\epsilon \leq x \leq \bar{y}^\epsilon \\ s_i^{F\epsilon}(x) & \text{if } \bar{y}^\epsilon < x < \bar{\bar{y}}^\epsilon \\ 0 & \text{if } x \geq \bar{\bar{y}}^\epsilon. \end{cases}$$

It can be easily checked that the output sharing $(\hat{s}_i^{F\epsilon})_{i\in T^F}$ always awards nonnegative payments to every agent in T^F, and also, the principal's residual is always nonnegative.

Proof of **Lemma 7**: Suppose $B(T^F) \geq V(T^*)$. Let $\epsilon > 0$ be such that $B(T^F) - |T^F|\epsilon > u_N$. Then it is obvious from (B1), (B2) and (B4) that

$$\text{(B5)} \qquad \hat{s}_i^{F\epsilon}(f|T^F(a_{T^F}^F)) - c_i(a_i^F) \;=\; u_i + \epsilon \quad \forall\ i \in T^F, \quad \text{and}$$

$$\text{(B6)} \qquad u_N + f|T^F(a_{T^F}^F) - \sum_{i\in T^F} \hat{s}_i^{F\epsilon}(f|T^F(a_{T^F}^F)) \;=\; B(T^F) - |T^F|\epsilon.$$

So it is sufficient to show that $a_{T^F}^F$ is the unique Nash equilibrium of $\{T^F, (\hat{s}_i^{F\epsilon})_{i\in T^F}\}$.

As $\underline{y}^\epsilon < u_i + \epsilon \ \forall\ i \in T^F$, using (B1) and (B3)-(B5), we can develop an argument similar to the one in the proof of Proposition 3 and show that $a_{T^F}^F$ is a Nash equilibrium of $\{T^F, (\hat{s}_i^{F\epsilon})_{i\in T^F}\}$. Thus, it remains to be shown that $\{T^F, (\hat{s}_i^{F\epsilon})_{i\in T^F}\}$ does not have any Nash equilibrium other than $a_{T^F}^F$.

(B4) and the limiting properties of the derivatives of c_i and f given in assumptions **A1** and **A2** imply that, if $a_{T^F} \in NE(\{T^F, (\hat{s}_i^{F\epsilon})_{i\in T^F}\})$ and $0 \leq f|T^F(a_{T^F}) < \underline{y}^\epsilon$, then $f_i|T^F(a_{T^F})/|T^F| - c_i'(a_i) = 0 \ \forall\ i \in T^F$. But it is obvious that, if $a_{T^F} \in A^{T^F}$ satisfies $f_i|T^F(a_{T^F})/|T^F| - c_i'(a_i) = 0 \ \forall\ i \in T^F$, then, because of the concavity of f and the strict convexity of c_i, a_{T^F} is a Nash equilibrium of $\{T^F, (s_i^e)_{i\in T^F}\}$, as given in Lemma 9. Hence, there does not exist $a_{T^F} \in NE(\{T^F, (\hat{s}_i^{F\epsilon})_{i\in T^F}\})$ such that $0 \leq f|T^F(a_{T^F}) < \underline{y}^\epsilon$.

Suppose $a_{T^F} \in NE(\{T^F, (\hat{s}_i^{F\epsilon})_{i\in T^F}\})$ is such that $\underline{y}^\epsilon \leq f|T^F(a_{T^F}) \leq \bar{y}^\epsilon$ or $f|T^F(a_{T^F}) \geq \bar{\bar{y}}^\epsilon$. Then there exists $j \in T^F$ such that $a_j > 0$, and hence, $c_j(a_j) > 0$. So $\hat{s}_j^{F\epsilon}(f|T^F(a_{T^F})) - c_j(a_j) = 0 - c_j(a_j) < 0 \leq \hat{s}_j^{F\epsilon}(f|T^F(a_{T^F_{-j}}, 0)) - c_j(0)$, where the last inequality follows from the fact that $\hat{s}_j^{F\epsilon}$ always pays a nonnegative amount to agent j and $c_j(0) = 0$. But this contradicts our supposition that $a_{T^F} \in NE(\{T^F, (\hat{s}_i^{F\epsilon})_{i\in T^F}\})$. Hence, there does not exist $a_{T^F} \in NE(\{T^F, (\hat{s}_i^{F\epsilon})_{i\in T^F}\})$ such that $\underline{y}^\epsilon \leq f|T^F(a_{T^F}) \leq \bar{y}^\epsilon$ or $f|T^F(a_{T^F}) \geq \bar{\bar{y}}^\epsilon$.

From (5) and (B4) it is clear that, if $a_{T^F} \in NE(\{T^F, (\hat{s}_i^{F\epsilon})_{i\in T^F}\})$ and $\bar{y}^\epsilon < f|T^F(a_{T^F}) < \bar{\bar{y}}^\epsilon$, then $f_i|T^F(a_{T^F}) - c_i'(a_i) = 0 \ \forall\ i \in T^F$. But we also know from the proof of (iv) of Proposition 1 that, if $a_{T^F} \in A^{T^F}$ and $f_i|T^F(a_{T^F}) - c_i'(a_i) = 0 \ \forall\ i \in T^F$, then $a_{T^F} = a_{T^F}^F$. So, if $a_{T^F} \in NE(\{T^F, (\hat{s}_i^{F\epsilon})_{i\in T^F}\})$ and $\bar{y}^\epsilon < f|T^F(a_{T^F}) < \bar{\bar{y}}^\epsilon$, then $a_{T^F} = a_{T^F}^F$.

So we have shown that $\{T^F, (\hat{s}_i^{F\epsilon})_{i\in T^F}\}$ has only one possible candidate for a Nash equilibrium, namely, $a^F_{T^F}$. Thus, as we already know that $a^F_{T^F} \in NE(\{T^F, (\hat{s}_i^{F\epsilon})_{i\in T^F}\})$, the proof of Lemma 7 is complete. ||

Suppose $V(T^*) > B(T^F)$. Then $V(T^*) > u_N \geq 0$. Now, pick any $\epsilon > 0$ such that $V(T^*) - |T^*_{-N}|\epsilon > u_N$ and $\alpha_i^* + \epsilon < 0 \; \forall \, i \in T^*_-$, and define $s_i^{*\epsilon}(x) = s_i^*(x) + \epsilon \; \forall \, x \in \Re_+$ and each $i \in T^*_{-N}$. Using Proposition 2 and the quasilinearity of the utility functions, it is obvious that

$$\text{(B7)} \qquad s_i^{*\epsilon}(f|T^*(a^*_{T^*})) - c_i(a_i^*) \;=\; u_i + \epsilon \quad \forall \, i \in T^*_{-N},$$

$$\text{(B8)} \qquad f|T^*(a^*_{T^*}) - \sum_{i\in T^*_{-N}} s_i^{*\epsilon}(f|T^*(a^*_{T^*})) - c_N(a_N^*) \;=\; V(T^*) - |T^*_{-N}|\epsilon, \quad \text{and}$$

$$\text{(B9)} \qquad NE(\{T^*, (s_i^{*\epsilon})_{i\in T^*_{-N}}\}) \;=\; \{a^*_{T^*}\}.$$

Let $y^{*\epsilon} > 0$ be such that $\beta_i^* y^{*\epsilon} < u_i + \epsilon \; \forall \, i \in T^*_{-N}$, and $(1 - \sum_{i\in T^*_{-N}} \beta_i^*) y^{*\epsilon} < V(T^*) - |T^*_{-N}|\epsilon$. Also, for each $i \in T^*_-$, let I_i^ϵ be the smallest positive integer such that $-(\alpha_i^* + \epsilon)/\beta_i^* \leq I_i^\epsilon y^{*\epsilon}$. Then, for each $i \in T^*_-$, define the payment function $\hat{s}_i^{*\epsilon}$ as follows:

$$\text{(B10)} \quad \hat{s}_i^{*\epsilon}(x) \;=\; \begin{cases} \beta_i^*[x - (I-1)y^{*\epsilon}] & \text{if } (I-1)y^{*\epsilon} \leq x < Iy^{*\epsilon} \text{ for } I = 1, ..., I_i^\epsilon \\ s_i^{*\epsilon}(x) & \text{if } x \geq I_i^\epsilon y^{*\epsilon}. \end{cases}$$

Given any nonnegative integer I and any $i \in T^*_-$, let

$$\lambda_i^\epsilon(I) \;=\; \begin{cases} 0 & \text{if } I = 0 \\ \beta_i^* y^{*\epsilon} & \text{if } 0 < I < I_i^\epsilon \\ \beta_i^* y^{*\epsilon} - s_i^{*\epsilon}(I_i^\epsilon y^{*\epsilon}) & \text{if } I = I_i^\epsilon \\ 0 & \text{if } I > I_i^\epsilon. \end{cases}$$

Then, for each nonnegative integer I, let $\lambda^\epsilon(I) = \sum_{i\in T^*_-} \lambda_i^\epsilon(I)$, and $\Lambda^\epsilon(I) = \sum_{\bar{I}=0}^{I} \lambda^\epsilon(\bar{I})$. Also, let $I^{\epsilon H}$ be the largest positive integer such that $\lambda^\epsilon(I^{\epsilon H}) > 0$.

Now, if $-\sum_{i\in T^*_{-N}} (\alpha_i^* + \epsilon) \geq 0$, then, for each $i \in T_0^* \cup T_+^*$, define the payment function $\hat{s}_i^{*\epsilon}$ as follows:

$$\text{(B11)} \quad \hat{s}_i^{*\epsilon}(x) = \begin{cases} \left[\dfrac{\alpha_i^*+\epsilon}{-\sum_{j\in T^*_-}(\alpha_j^*+\epsilon)}\right] \Lambda^\epsilon(I) + \beta_i^* x & \text{if } Iy^{*\epsilon} \leq x < (I+1)y^{*\epsilon} \text{ for } I = 0, ..., I^{\epsilon H} - 1 \\ s_i^{*\epsilon}(x) & \text{if } x \geq I^{\epsilon H} y^{*\epsilon}. \end{cases}$$

On the otherhand, if $-\sum_{i\in T^*_{-N}} (\alpha_i^* + \epsilon) < 0$, then let I_N^ϵ be the smallest positive integer such that $\sum_{i\in T^*_{-N}} (\alpha_i^* + \epsilon)/(1 - \sum_{i\in T^*_{-N}} \beta_i^*) \leq I_N^\epsilon y^{*\epsilon}$. Now, given any nonnegative integer I, define

$$\lambda_N^\epsilon(I) \;=\; \begin{cases} 0 & \text{if } I = 0 \\ (1 - \sum_{i\in T^*_{-N}} \beta_i^*) y^{*\epsilon} & \text{if } 0 < I < I_N^\epsilon \\ (1 - \sum_{i\in T^*_{-N}} \beta_i^*) y^{*\epsilon} - (x - \sum_{i\in T^*_{-N}} s_i^{*\epsilon}(I_N^\epsilon y^{*\epsilon})) & \text{if } I = I_N^\epsilon \\ 0 & \text{if } I > I_N^\epsilon. \end{cases}$$

Then, for each nonnegative integer I, let $\bar{\lambda}^\epsilon(I) = \lambda^\epsilon(I) + \lambda_N^\epsilon(I)$, and $\bar{\Lambda}^\epsilon(I) = \sum_{\bar{I}=0}^{I} \bar{\lambda}^\epsilon(\bar{I})$. Also, let $\bar{I}^{\epsilon H}$ be the largest positive integer such that $\bar{\lambda}^\epsilon(I^{\epsilon H}) > 0$.

Thus, if $-\sum_{i\in T^*_{-N}} (\alpha_i^* + \epsilon) < 0$, then, for each $i \in T_0^* \cup T_+^*$, replace the definition of $\hat{s}_i^{*\epsilon}$ in (B11) by the following:

$$\text{(B12)} \quad \hat{s}_i^{*\epsilon}(x) = \begin{cases} \left[\dfrac{\alpha_i^*+\epsilon}{\sum_{j\in T_0^*\cup T_+^*}(\alpha_j^*+\epsilon)}\right] \bar{\Lambda}^\epsilon(I) + \beta_i^* x & \text{if } Iy^{*\epsilon} \leq x < (I+1)y^{*\epsilon} \text{ for } I = 0, ..., \bar{I}^{\epsilon H} - 1 \\ s_i^{*\epsilon}(x) & \text{if } x \geq \bar{I}^{\epsilon H} y^{*\epsilon}. \end{cases}$$

It can be verified that the output sharing $(\hat{s}_i^{*\epsilon})_{i\in T^*_{-N}}$, as defined above, always awards nonnegative payments to every agent in T^*_{-N}, and also, the principal's residual is always nonnegative.

Proof of **Lemma 8**: Suppose $V(T^*) > B(T^F)$. Let $\epsilon > 0$ be such that $V(T^*) - |T^*_{-N}|\epsilon > u_N$ and $\alpha_i^* + \epsilon < 0$ $\forall\ i \in T^*_-$. Then, using (B7), (B8) and (B10)-(B12), it can be checked that

$$\text{(B13)} \qquad \hat{s}_i^{*\epsilon}(f|T^*(a^*_{T^*})) - c_i(a_i^*) \;=\; u_i + \epsilon \quad \forall\ i \in T^*_{-N}, \text{ and}$$

$$\text{(B14)} \quad f|T^*(a^*_{T^*}) - \sum_{i\in T^*_{-N}} \hat{s}_i^{*\epsilon}(f|T^*(a^*_{T^*})) - c_N(a_N^*) \;=\; V(T^*) - |T^*_{-N}|\epsilon.$$

So it is sufficient to show that $a^*_{T^*}$ is the unique Nash equilibrium of $\{T^*, (\hat{s}_i^{*\epsilon})_{i\in T^*_{-N}}\}$.

For each $i \in T^*_-$, if $x \in [0, I_i^\epsilon y^{*\epsilon})$, then (B10) implies that $\hat{s}_i^{*\epsilon}(x) < u_i + \epsilon$. So, for each $i \in T^*_-$, as c_i is strictly increasing and $c_i(0) = 0$, (B10) and (B13) imply that $\hat{s}_i^{*\epsilon}(f|T^*(a^*_{T^*_{-i}}, a_i)) - c_i(a_i) \leq \hat{s}_i^{*\epsilon}(f|T^*(a^*_{T^*})) - c_i(a_i^*)$ for any $a_i \in A_i$ such that $0 \leq f|T^*(a^*_{T^*_{-i}}, a_i) < I_i^\epsilon y^{*\epsilon}$.

Also, it is clear from (B9) and (B10) that, for each $i \in T^*_-$, $\hat{s}_i^{*\epsilon}(f|T^*(a^*_{T^*_{-i}}, a_i)) - c_i(a_i) \leq \hat{s}_i^{*\epsilon}(f|T^*(a^*_{T^*})) - c_i(a_i^*)$ for any $a_i \in A_i$ such that $f|T^*(a^*_{T^*_{-i}}, a_i) \geq I_i^\epsilon y^{*\epsilon}$.

For each $i \in T_0^* \cup T_+^*$, using (B11) or (B12) depending on whether $-\sum_{j\in T^*_{-N}}(\alpha_j^* + \epsilon)$ is nonnegative or not, it can be checked that $\hat{s}_i^{*\epsilon}(x) \leq s_i^{*\epsilon}(x)\ \forall\ x \in \Re_+$. Thus, for each $i \in T_0^* \cup T_+^*$, because of (B7), (B9) and (B13), $\hat{s}_i^{*\epsilon}(f|T^*(a^*_{T^*_{-i}}, a_i)) - c_i(a_i) \leq \hat{s}_i^{*\epsilon}(f|T^*(a^*_{T^*})) - c_i(a_i^*)\ \forall\ a_i \in A_i$.

It can be verified that, if $-\sum_{j\in T^*_{-N}}(\alpha_j^* + \epsilon) \geq 0$, then (B10) and (B11) imply $x - \sum_{j\in T^*_{-N}} \hat{s}_j^{*\epsilon}(x) \leq x - \sum_{j\in T^*_{-N}} s_j^{*\epsilon}(x)\ \forall\ x \in \Re_+$. Thus, because of (B8), (B9) and (B14), if $-\sum_{j\in T^*_{-N}}(\alpha_j^* + \epsilon) \geq 0$, then $f|T^*(a^*_{T^*_{-N}}, a_N) - \sum_{j\in T^*_{-N}} \hat{s}_j^{*\epsilon}(f|T^*(a^*_{T^*_{-N}}, a_N)) - c_N(a_N) \leq f|T^*(a^*_{T^*}) - \sum_{j\in T^*_{-N}} \hat{s}_j^{*\epsilon}(f|T^*(a^*_{T^*})) - c_N(a_N^*)\ \forall\ a_N \in A_N$.

It can also be verified that, if $-\sum_{j\in T^*_{-N}}(\alpha_j^* + \epsilon) < 0$, then (B10) and (B12) imply $x - \sum_{j\in T^*_{-N}} \hat{s}_j^{*\epsilon}(x) < V(T^*) - |T^*_{-N}|\epsilon\ \forall\ x \in [0, I_N^\epsilon y^{*\epsilon})$. So, if $-\sum_{j\in T^*_{-N}}(\alpha_j^* + \epsilon) < 0$, then the monotonicity of c_N, $c_N(0) = 0$ and (B14) imply that $f|T^*(a^*_{T^*_{-N}}, a_N) - \sum_{j\in T^*_{-N}} \hat{s}_j^{*\epsilon}(f|T^*(a^*_{T^*_{-N}}, a_N)) - c_N(a_N) \leq f|T^*(a^*_{T^*}) - \sum_{j\in T^*_{-N}} \hat{s}_j^{*\epsilon}(f|T^*(a^*_{T^*})) - c_N(a_N^*)$ for any $a_N \in A_N$ such that $0 \leq f|T^*(a^*_{T^*_{-N}}, a_N) < I_N^\epsilon y^{*\epsilon}$.

Once again, it can be verified that, if $-\sum_{j\in T^*_{-N}}(\alpha_j^* + \epsilon) < 0$, then (B10) and (B12) imply $x - \sum_{j\in T^*_{-N}} \hat{s}_j^{*\epsilon}(x) = x - \sum_{j\in T^*_{-N}} s_j^{*\epsilon}(x)\ \forall\ x \geq I_N^\epsilon y^{*\epsilon}$. So, if $-\sum_{j\in T^*_{-N}}(\alpha_j^* + \epsilon) < 0$, then, because of (B9), $f|T^*(a^*_{T^*_{-N}}, a_N) - \sum_{j\in T^*_{-N}} \hat{s}_j^{*\epsilon}(f|T^*(a^*_{T^*_{-N}}, a_N)) - c_N(a_N) \leq f|T^*(a^*_{T^*}) - \sum_{j\in T^*_{-N}} \hat{s}_j^{*\epsilon}(f|T^*(a^*_{T^*})) - c_N(a_N^*)$ for any $a_N \in A_N$ such that $f|T^*(a^*_{T^*_{-N}}, a_N) \geq I_N^\epsilon y^{*\epsilon}$.

Thus, we have shown that $a^*_{T^*} \in NE(\{T^*, (\hat{s}_j^{*\epsilon})_{j\in T^*_{-N}}\})$.

For each $i \in T^*_{-N}$, it is clear from the definitions in (B10)-(B12) that $\hat{s}_i^{*\epsilon}$ is piecewise linear and has a slope of β_i^* (> 0) on each linear piece. Similarly, (B10)-(B12) also imply that $x - \sum_{j\in T^*_{-N}} \hat{s}_j^{*\epsilon}(x)$, the principal's residual, is piecewise linear in $x \in \Re_+$ and has a slope of $1 - \sum_{j\in T^*_{-N}} \beta_j^*$ (> 0) on each linear piece. Therefore, the limiting properties of the derivatives of the cost functions and the production function in assumptions **A1** and **A2** immediately imply that, if $a_{T^*} \in NE(\{T^*, (\hat{s}_j^{*\epsilon})_{j\in T^*_{-N}}\})$, then $a_i > 0\ \forall\ i \in T^*$.

It can be easily checked that the proof of Lemma 3 uses only the Nash equilibrium condition in constraint (C2). So, if $a_{T^*} \in NE(\{T^*, (\hat{s}_j^{*\epsilon})_{j\in T^*_{-N}}\})$, then an argument similar to the one in the proof of Lemma 3 shows that, for each $i \in T^*_{-N}$, $\hat{s}_i^{*\epsilon}$ is differentiable at $f|T^*(a_{T^*})$. Hence, if $a_{T^*} \in NE(\{T^*, (\hat{s}_j^{*\epsilon})_{j\in T^*_{-N}}\})$, then (B10)-(B12) imply that $\beta_i^* = c_i'(a_i)/f_i|T^*(a_{T^*})\ \forall\ i \in T^*_{-N}$ and $1 - \sum_{j\in T^*_{-N}} \beta_j^* = c_N'(a_N)/f_N|T^*(a_{T^*})$. However, we know from the proof of (iv) of Proposition 2 that $a_{T^*} \in A^{T^*}$ satisfies $\beta_i^* = c_i'(a_i)/f_i|T^*(a_{T^*})\ \forall\ i \in T^*_{-N}$ and $1 - \sum_{j\in T^*_{-N}} \beta_j^* = c_N'(a_N)/f_N|T^*(a_{T^*})$ if and only if $a_{T^*} = a^*_{T^*}$. Therefore, as we have already shown that $a^*_{T^*} \in NE(\{T^*, (\hat{s}_j^{*\epsilon})_{j\in T^*_{-N}}\})$, the proof of Lemma 8 is complete. ||

References

[1] Alchian, A. and Demsetz,H. "Production, Information Costs, and Economic Organization", *American Economic Review.* 62 (1972), 777–795.

[2] Arrow, K. *Limits of Organization.* New York: Norton & Co. (1974).

[3] Avriel, M., Diewert, W.E., Schaible, S. and Zang, I. *Generalized Concavity.* Plenum Press, New York (1988).

[4] Carmichael, H. L. "The Agent-Agents Problem: Payment by Relative output", *Journal of Labour Economics.* 1 (1983), 50–65.

[5] Eswaran, M. and Kotwal, A. "The Moral Hazard of Budget-breaking", *Rand Journal of Economics.* 15 (1984), 578–581.

[6] ————. "A Theory of Contractual Structure in Agriculture", *American Economic Review.* 75 (1985), 352–367.

[7] Grossman, S. and Hart, O. "An Analysis of the Principal-Agent Problem", *Econometrica.* 51 (1983), 7–45.

[8] Harris, M. and Raviv, A. "Optimal Incentive Contracts with Imperfect Information", *Journal of Economic Theory.* 20 (1979), 231–259.

[9] Hart, O. and Holmstrom, B. "The Theory of Contracts", in Bewley, T. F., ed., *Advances in Economic Theory: Fifth World Congress.* Cambridge University Press (1987).

[10] Holmstrom, B. "Moral Hazard and Observability", *Bell Journal of Economics.* 10 (1979), 74–91.

[11] ————. "Moral Hazard in Teams", *Bell Journal of Economics.* 13 (1982), 324–340.

[12] Legros, P. and Matsushima, H. "Efficiency in Partnerships", *Journal of Economic Theory.* 55 (1991), 296–322.

[13] Mirrlees, J. "On the Optimal Structure of Incentives, and Control within an Organization", *Bell Journal of Economics.* 7 (1976), 105–131.

[14] Rasmusen, E. "Moral Hazard in Risk-Averse Teams", *Rand Journal of Economics.* 18 (1987), 428–435.

[15] Rees, R. "The Theory of Principal and Agent: Part I", *Bulletin of Economic Research.* 37 (1985), 3–26.

[16] Ross, S. "The Economic Theory of Agency: The Principal's Problem", *American Economic Review* (Paper and Proceedings). 63 (1973), 134–139.

[17] Selten, R. "Re-examination of the Perfectness Concept for Equilibrium Points in Extensive Games", *Int. J. Game Theory.* 4 (1975), 25–55.

[18] Shavell, S. "Risk Sharing and Incentives in the Principal and Agent Relationship", *Bell Journal of Economics.* 10 (1979), 55–73.

[19] Stiglitz, J. "Incentives and Risk Sharing in Sharecropping", *Review of Economic Studies.* 61 (1974), 219–256.

[20] Williams, S. and Radner, R. "Efficiency in Partnerships when the Joint Output is Uncertain", *Discussion Paper No.* 760 (1988), Northwestern University.

[21] Wilson, R. "The Theory of Syndicates", *Econometrica.* 36 (1968), 119–132.

CENTRE FOR DEVELOPMENT ECONOMICS WORKING PAPER SERIES

Number	Author(s)	Title
1	Kaushik Basu Arghya Ghosh Tridip Ray	The Babu and The Boxwallah : Managerial Incentives and Government Intervention (Jan 1994)
2	M. N. Murty Ranjan Ray	Optimal Taxation and Resource Transfers in a Federal Nation (Feb 1994)
3	V. Bhaskar Mushtaq Khan	Privatization and Employment : A Study of The Jute Industry in Bangladesh (Mar 1994)
4	V. Bhaskar	Distributive Justice and The Control of Global Warming (Mar 1994)
5	Bishnupriya Gupta	The Great Depression and Brazil's Capital Goods Sector : A Re-examination (Apr 1994)
6.	Kaushik Basu	Where There Is No Economist : Some Institutional and Legal Prerequisites of Economic Reform in India (May 1994)
7.	Partha Sen	An Example of Welfare Reducing Tariff Under Monopolistic Competition (May 1994)
8.	Partha Sen	Environmental Policies and North-South Trade : A Selected Survey of The Issues (May 1994)
9.	Partha Sen Arghya Ghosh Abheek Barman	The Possibility of Welfare Gains with Capital Inflows in A Small Tariff-Ridden Economy (June 1994)
10.	V. Bhaskar	Sustaining Inter-Generational Altruism when Social Memory is Bounded (June 1994)
11.	V. Bhaskar	Repeated Games with Almost Perfect Monitoring by Privately Observed Signals (June 1994)

Number	Author(s)	Title
12.	S. Nandeibam	Coalitional Power Structure in Stochastic Social Choice Functions with An Unrestricted Preference Domain (June 1994)
13.	Kaushik Basu	The Axiomatic Structure of Knowledge And Perception (July 1994)
14.	Kaushik Basu	Bargaining with Set-Valued Disagreement (July 1994)
15.	S. Nandeibam	A Note on Randomized Social Choice and Random Dictatorships (July 1994)
16.	Mrinal Datta Chaudhuri	Labour Markets As Social Institutions in India (July 1994)
17.	S. Nandeibam	Moral Hazard in a Principal-Agent(s) Team (July 1994)