INTERNATIONAL ECONOMIC REVIEW
Vol. 41, No. 2, May 2000

SURPLUS SHARING WITH A TWO-STAGE MECHANISM*

By Todd R. Kaplan and David Wettstein[1]

Ben-Gurion University of the Negev, Israel

In this article we consider environments where agents jointly produce a private output good by contributing privately owned resources. An efficient outcome may not be realized due to strategic behavior and conflicting interests of the agents. We construct a two-stage mechanism, building on a Varian mechanism. The modified mechanism ensures an equilibrium for a large class of preferences and guarantees the feasibility of outcomes.

1. INTRODUCTION

Achieving efficiency in environments characterized by joint ownership of a technology is not straightforward. Agents may have incentive to overutilize the technology, such as in the "tragedy of the commons," or underutilize the technology, as with the free-rider problem. We focus on environments where the technology output is a private good that leads to what has commonly been referred to as *surplus-sharing problems* (with an analogous situation called the *cost-sharing problem*).

Several mechanisms have been proposed to solve this problem: markets, average-cost pricing, marginal-cost pricing, and serial-cost sharing [see Young (1994) for a survey]. The choice of either markets or marginal-cost pricing is an inadequate solution due to possible strategic behavior on part of the agents (see Hurwicz, 1972), while average-cost pricing (Moulin and Watts, 1997) and serial-cost sharing (Moulin and Shenker, 1992) are not designed to guarantee efficiency. Kaplan and Wettstein (1996) constructed a mechanism that imitates to a large degree the operation of markets and achieves an efficient outcome. Market-type mechanisms that achieve efficient outcomes for related environments with public goods also have been suggested by Tian (1994).

Another mechanism proposed by Roemer and Silvestre (1993) is to divide the output of the technology in proportion to the individuals' input contributions. An efficient outcome resulting from this mechanism is called a *proportional solution*. Moulin (1990) provides an axiomatic foundation for the proportional solution, and Roemer and Silvestre (1993) prove its existence for convex environments. Strategic behavior might once more undermine efficiency; however, Suh (1995) constructs a

* Manuscript received April 1997; revised August 1998.

[1] We thank two anonymous referees for their insightful comments that substantially improved this article. We gratefully acknowledge the support of the Kreitman Foundation and the Monaster Center for Economic Research.

mechanism whose Nash equilibria (as well as strong Nash and undominated Nash equilibria) outcomes coincide with the proportional solution outcomes.

We construct a two-stage mechanism in the spirit of implementation theory [see Moore (1992) for a recent survey and Hurwicz (1960)] by stressing both simplicity and information requirements. It is simpler than the Kaplan and Wettstein (1996) mechanism due to the use of two stages and simpler than the general Moore and Repullo (1988) multistage mechanism due to the specific structure of the problem we analyze. It also has less stringent informational requirements than the Suh (1995) mechanism while providing several different sets of efficient outcomes. Each efficient outcome realized depends on the designer's choice of interim output shares. For instance, choosing equal shares of output, common practice on a kibbutz, leads to one particular efficient outcome.

Saijo et al. (1996) propose properties that would be desired in a "natural" mechanism and then analyze the implications of these properties for Nash implementation in pure-exchange economies. Our mechanism enjoys similar properties but uses subgame perfection in a two-stage game for a production environment.

The mechanism designed is an extension of a two-stage mechanism constructed in Varian (1994). Varian's goal was to realize efficient outcomes in environments that involve externalities. The externalities present in our environment stem from the fact that the input rendered by one agent affects the productivity of the input contributed by another; the output requested by one person affects the cost of providing the output demanded by another person. These externalities are actually present in any competitive environment with decreasing returns to scale. We extend Varian's construction to work in our environment for a large class of preferences. Our extension provides feasible outcomes both in and out of equilibrium. We also show that the added complexity does not interfere with the existence of equilibria.

We describe and define the environment and the mechanism in Sections 2 and 3. In Section 4 we prove that subgame-perfect equilibria exist and are optimal. Conclusions and further directions of research are offered in the final section.

2. ENVIRONMENT

We consider a surplus-sharing problem where there are three agents that consume two goods, x and y. Agents are endowed with w_i (> 0) of good x and none of good y. Good y can be produced via the production function $f(\cdot)$, which is assumed to be strictly increasing, concave, and differentiable.[2] Agent i's preferences can be represented by a utility function u^i that is strictly increasing, differentiable, concave, and satisfies the Inada conditions $[\lim_{x_i \to 0} u_1^i(x_i, y_i) = \lim_{y_i \to 0} u_2^i(x_i, y_i) = \infty, \lim_{x_i \to \infty} u_1^i(x_i, y_i) = \lim_{y_i \to \infty} u_2^i(x_i, y_i) = 0]$.

[2] Since the technology is shared, an agent is unable to produce on his or her own. Note that if the equilibrium outcome of a mechanism gives strictly positive consumption levels, then it is individually rational for the agents to take part in that mechanism.

An allocation consisting of the agents' consumption levels (x_i, y_i) and input-output levels x_p, y_p is feasible if

$$x_p + \sum_{i=1}^{3} x_i \leq \sum_{i=1}^{3} w_i \tag{1}$$

$$\sum_{i=1}^{3} y_i \leq y_p \tag{2}$$

$$y_p \leq f(x_p) \tag{3}$$

Notice that in a feasible allocation there may be an agent i that consumes more of good x than his or her initial endowment $(x_i > w_i)$.

The first-order conditions characterizing a Pareto-optimal allocation can be derived in a straightforward manner and are given by $u_1^1/u_2^1 = u_1^2/u_2^2 = u_1^3/u_2^3 = f'$.

Each agent is fully informed about the environment (including explicit knowledge of f); however, the designer is not. He or she knows the endowments and is able to observe input and output levels but knows neither the technology nor the agents' preferences. In what follows we describe a mechanism where the agents send messages to the designer, who in turn decides on production and consumption plans. We will later show that the equilibria of the mechanism give rise to efficient outcomes.

3. MECHANISM

The mechanism consists of two stages. In the first stage, the agents simultaneously announce prices. These prices determine how much agents must pay for the consumption of good x. In the second stage, the agents choose suggested consumption levels of good x. If the sum of suggested consumptions is less than the total aggregate endowment, the difference is used as the production input. The mechanism then allocates the production output by solving a system of equations imitating budget constraints that incorporate the prices from the first stage. If either the sum of announced x's exceeds total endowment or the equations yield negative consumption levels for one or more of the agents (violating feasibility), the mechanism goes through a "punishment phase."

We will now formally describe the mechanism. In stage 1, the prices announced by the agents are (p_{jk}^i, p_{kj}^i) (both prices must be strictly positive), where p_{jk}^i refers to a price chosen by agent i that agent k must pay agent j for each unit of good x. All prices are revealed to everyone before stage 2. In stage 2, agents simultaneously announce x_i's, which must be strictly positive.

If the sum of the x_i's is greater than the sum of endowments W, the mechanism goes into a punishment phase.[3] It finds any agents that could have prevented the sum of x_i's from being greater than W and gives them $(w_i, 0)$. If no such agents are found, then it gives $(w_i, 0)$ to all agents submitting x_i's (weakly) larger than w_i. The mechanism then determines the set of agents I that have not thus far been allocated

[3] Dutta et al. (1995) and Saijo et al. (1996) employ similar punishment schemes.

$(w_i, 0)$. If the number of agents in I ($\#I$) is greater than zero, it gives those agents $(w_i/2, f(\sum_{i\in I} w_i/2)/\#I)$. Even though f explicitly appears in the description of the mechanism, the designer does not need to explicitly know f; all that is required is the ability he or she has to engage the technology using a specific amount of input and to receive the output that is produced.

Formally defined, in the case where $x_1 + x_2 + x_3 \geq W$, let $S_1 = \{i \mid \sum_{j\neq i} x_j < W\}$ and $S_2 = \{i \mid x_i \geq w_i\}$, and let the bundle agent i receives be given by

$$h_i(x_1, x_2, x_3) = \begin{cases} (w_i, 0) & \text{if } i \in S_1 \\ \left(w_i/2, \dfrac{f(\sum_{i\notin S_1} w_i/2)}{n - \#S_1}\right) & \text{if } 0 < \#S_1 < n \text{ and } i \notin S_1 \\ (w_i, 0) & \text{if } S_1 = \Phi \text{ and } i \in S_2 \\ \left(w_i/2, \dfrac{f(\sum_{i\notin S_2} w_i/2)}{n - \#S_2}\right) & \text{if } S_1 = \Phi, \#S_2 < n \text{ and } i \notin S_2 \end{cases}$$

If the x_i's pass the initial test, the mechanism produces output using an input level x_p that is equal to $\sum_{i=1}^{3}(w_i - x_i)$ and sets up the following system of equations:

$$\begin{aligned} y_1 &= \alpha_1 \cdot f(x_p) - (p_{31}^2 + p_{21}^3) \cdot x_1 + p_{12}^3 \cdot x_2 + p_{13}^2 \cdot x_3 \\ y_2 &= \alpha_2 \cdot f(x_p) - (p_{32}^1 + p_{12}^3) \cdot x_2 + p_{21}^3 \cdot x_1 + p_{23}^1 \cdot x_3 \\ y_3 &= \alpha_3 \cdot f(x_p) - (p_{23}^1 + p_{13}^2) \cdot x_3 + p_{31}^2 \cdot x_1 + p_{32}^1 \cdot x_2 \end{aligned} \tag{4}$$

The α_i's in the equations must be strictly positive and sum to one. The α_i's can be interpreted as interim output shares. Note that a choice of different α_i's corresponds to a different mechanism and thus possibly leads to a different allocation and that the final y_i's allocated usually would be different from the interim share of output $\alpha_i \cdot f(x_p)$.

If the solution of these equations yields nonnegative y_i's, the mechanism ends with the consumption bundle allocated to agent i being (x_i, y_i). If the solution entails at least one negative y_i, the mechanism allocates $(0, 0)$ to all those with negative y_i's in the preceding equations and assigns allocations $(x_i, \max\{\widehat{y}_i, 0\})$ to all others, where

$$\begin{aligned} \widehat{y}_1 &= \alpha_1 \cdot f(x_p) - (p_{31}^2 + p_{21}^3) \cdot x_1 \\ \widehat{y}_2 &= \alpha_2 \cdot f(x_p) - (p_{32}^1 + p_{12}^3) \cdot x_2 \\ \widehat{y}_3 &= \alpha_3 \cdot f(x_p) - (p_{23}^1 + p_{13}^2) \cdot x_3 \end{aligned}$$

The mechanism uses finite-dimensional message spaces consisting of quantity and price announcements. Furthermore, this mechanism is feasible for any set of messages. The mechanism always allocates nonnegative consumption. The first equation (1) of feasibility is guaranteed by the first punishment phase. The second and third equations (2 and 3) of feasibility are satisfied for nonnegative y_i's because adding together the equations in (4) yields $\sum y_i = f(x_p)$. When one of the y_i's turns out negative, the sum of the $\widehat{y}_i$'s does not exceed $f(x_p)$, and hence the feasibility conditions hold under this contingency.

A "natural" mechanism as defined in Saijo et al. (1996) has a finite-dimensional strategy space consisting of price and quantity announcements, forthrightness, both individual and aggregate feasibility (nonnegative consumption and balanced outcomes) for any strategy profile, and a best-response property. Forthrightness demands that when the price and quantity announcements by individual players are consistent with a socially desired allocation, these announcements form an equilibrium yielding that allocation. The best-response property entails that for every possible strategy profile, a best response exists for each individual player.

When these properties are adjusted to a two-stage game, our mechanism satisfies most of them. The strategy spaces at each stage are just price and quantity announcements. When the strategy choices in both stages coincide with the quantity and shadow prices of an efficient outcome, this will be a subgame-perfect equilibrium yielding the efficient outcome. The best-response property is satisfied in the second stage of the game. It is not satisfied in the first stage because discontinuous behavior (out of equilibrium) by one player in the second stage (quantity being a discontinuous function of price) could make it impossible for another player to respond optimally in the first stage. Finally, the outcome function is individually feasible by construction but is only weakly balanced for some cases of punishment (we assume free disposal).

4. OPTIMALITY AND EXISTENCE OF SUBGAME-PERFECT ALLOCATIONS

We use the notion of subgame perfection as our equilibrium concept. The strategy of an agent can be described by a choice of prices and a choice of x_i for each possible set of prices. Together, the set of strategies form a subgame-perfect equilibrium if for each possible set of prices, the choices of x_i's are a Nash equilibrium and, given the contingent future choices of x_i's, the choice of prices is a Nash equilibrium. Initially, we show that all subgame-perfect equilibria are Pareto optimal. Afterwards, we show that subgame-perfect equilibria exist. To demonstrate optimality, we show that equilibria allocations must be interior, and then the agents' choices of both prices and quantities are shown to result in first-order conditions that satisfy Pareto optimality. The intuition is that we create a game where each agent i receives an α_i fraction of the output independent of how much input he or she contributes. The externality agent i imposes on agent j by consuming more input good is the reduction of the α_j share of output. This externality is internalized because in equilibrium the price he or she pays to each agent for consuming input goods is equal to the marginal externality he or she imposes on him or her. These prices are chosen correctly because each agent only chooses prices that other agents must pay each other.[4]

LEMMA 1. *A subgame perfect equilibrium path of the mechanism does not involve any punishment.*

[4] The construction whereby parts of an individual's message affect only other individuals and furthermore each individual is affected exactly by two other individuals is quite common in the implementation literature; see Walker (1981) for an early contribution along those lines.

PROOF. There are two occasions where punishment may occur. Any punishment involves giving at least one agent $(w_i, 0)$. If punishment is invoked because the sum of x_i's exceeds aggregate endowments, then any agent receiving $(w_i, 0)$ can improve his or her allocation because he or she could either move the game out of this punishment phase or announce an x_i strictly less than w_i. If punishment is imposed due to a negative y_i, then any agent who receives $(0, 0)$ can improve his or her allocation by announcing an x_i that is arbitrarily close to 0. ■

LEMMA 2. *Any subgame-perfect equilibrium allocation is interior.*

PROOF. Since there is no punishment in equilibrium, the consumption of the x good is always strictly positive. If agent i receives a zero amount of the y good, he or she can announce a small enough x_i that will give him or her strictly positive consumption of both goods. Even if a punishment phase is invoked by decreasing his or her x_i (due to another agent's y_j turning negative), it is always possible for agent i to get strictly positive amounts of both commodities. ■

LEMMA 3. *All subgame-perfect equilibrium allocations satisfy the following equations*:

$$u_1^1 = u_2^1\left[\alpha_1 f'(x_p) + p_{31}^2 + p_{21}^3\right]$$
$$u_1^2 = u_2^2\left[\alpha_2 f'(x_p) + p_{32}^1 + p_{12}^3\right]$$
$$u_1^3 = u_2^3\left[\alpha_3 f'(x_p) + p_{23}^1 + p_{13}^2\right]$$
$$\alpha_1 f'(x_p) = p_{12}^3 = p_{13}^2$$
$$\alpha_2 f'(x_p) = p_{21}^3 = p_{23}^1$$
$$\alpha_3 f'(x_p) = p_{31}^2 = p_{32}^1$$

PROOF. Denote by $x_i(p_{23}^1, p_{32}^1, p_{13}^2, p_{31}^2, p_{12}^3, p_{21}^3)$ the quantities demanded by agent i at the second stage of the game. The problem facing agent 1 in the first stage is to choose an optimal pair of prices p_{23}^1, p_{32}^1:

$$\max_{p_{23}^1, p_{32}^1} u^1\left\{x_1(p_{23}^1, p_{32}^1, \ldots), \alpha_1 f\left[\sum w_i - x_i(p_{23}^1, p_{32}^1, \ldots)\right]\right.$$
$$\left. - (p_{31}^2 + p_{21}^3)x_1(p_{23}^1, p_{32}^1, \ldots) + p_{12}^3 x_2(p_{23}^1, p_{32}^1, \ldots) + p_{13}^2 x_3(p_{23}^1, p_{32}^1, \ldots)\right\}$$

Differentiating with respect to p_{23}^1 and p_{32}^1 yields the following two equations:

$$[u_1^1 - u_2^1(\alpha_1 f' + p_{31}^2 + p_{21}^3)]\frac{\partial x_1}{\partial p_{23}^1} + (p_{12}^3 - \alpha_1 f')\frac{\partial x_2}{\partial p_{23}^1} + (p_{13}^2 - \alpha_1 f')\frac{\partial x_3}{\partial p_{23}^1} = 0$$
$$[u_1^1 - u_2^1(\alpha_1 f' + p_{31}^2 + p_{21}^3)]\frac{\partial x_1}{\partial p_{32}^1} + (p_{12}^3 - \alpha_1 f')\frac{\partial x_2}{\partial p_{32}^1} + (p_{13}^2 - \alpha_1 f')\frac{\partial x_3}{\partial p_{32}^1} = 0$$

In the second stage, agent 1 chooses an optimal x_1 while taking the prices as given. This leads to the equation $u_1^1 - u_2^1(\alpha_1 f' + p_{31}^2 + p_{21}^3) = 0$.

We assume that the matrix

$$\begin{bmatrix} \dfrac{\partial x_2}{\partial p_{23}^1} & \dfrac{\partial x_3}{\partial p_{23}^1} \\ \dfrac{\partial x_2}{\partial p_{32}^1} & \dfrac{\partial x_3}{\partial p_{32}^1} \end{bmatrix}$$

is nonsingular.[5] Hence, for these three equations to be satisfied, $\alpha_1 f' = p_{12}^3$ and $\alpha_1 f' = p_{13}^2$.

By considering the problems facing agents 2 and 3, we can show the rest of the conditions of this lemma.

Intuitively, the first set of equations is the set of first-order conditions for the choices of x_i's in stage 2. The second set of equations is derived from the agents' choice of p_{ij}^k. The choice of prices in the first stage determines the subgame that will be played. This indirectly determines the x_i's that will be chosen. Given the nonsingularity of the matrix described earlier, agent i is able to move the other agents' choices in any direction. Since an agent's choice of prices has no direct influence on his or her budget set and an agent would like to have the other agents choose their x_i's in such a way as to make his or her budget set the largest, it must be that $\alpha_i f'(x_p) = p_{ij}^k$ for $(i \neq j)$. ■

THEOREM 1. *All subgame-perfect equilibrium allocations are Pareto optimal.*

PROOF. Combining the two sets of equations from Lemma 3 yields $u_1^1/u_2^1 = u_1^2/u_2^2 = u_1^3/u_2^3 = f'$. These conditions assure us of Pareto optimality because preferences are convex and the production function is concave. ■

So far we have demonstrated that subgame-perfect equilibrium allocations are Pareto optimal; to show that this result is not vacuously satisfied, we prove that there are Pareto-optimal allocations that can be realized as a subgame-perfect equilibrium allocation. We will show this by constructing an auxiliary economy where competitive equilibria exist and then proving that these have analogous subgame-perfect equilibria in our mechanism.

Take a competitive economy with the same preferences and technology as our environment where the agents own the firm with shares given by α_i's and have initial endowments $\overline{w}_i = \alpha_i \sum_{i=1}^N w_i$. The budget constraint of agent i is

$$x_i + q \cdot y_i \leq \overline{w}_i + \alpha_i \left[q \cdot f(x_p) - x_p \right]$$

where q is the price of good y in terms of good x, and x_p is the production input. Given our assumptions on preferences and technology, an interior competitive equilibrium exists and gives rise to a Pareto-optimal allocation. We will now use the analogous equilibria of our mechanism to prove the following:

THEOREM 2. *There exists a subgame-perfect equilibrium for our mechanism.*

[5] We thank an annonymous referee for pointing out the necessity of making this assumption.

PROOF. Construct a competitive economy as described earlier, and consider an allocation $\{(x'_i, y'_i)_{i=1}^3, x'_p\}$ generated by an interior competitive equilibrium. Since the firm is maximizing profits, $q = 1/f'(x'_p)$. Since agents are maximizing their utility, the budget constraints are satisfied with equality:

$$x'_i + q \cdot y'_i = \overline{w}_i + \alpha_i \left[q \cdot f(x'_p) - x'_p\right] \tag{5}$$

We will now define a set of strategies and show they constitute a subgame-perfect equilibrium. The first-stage strategies will be

$$\begin{aligned} \widetilde{p}^1_{32} &= \alpha_3 f'(x'_p);\ \widetilde{p}^1_{23} = \alpha_2 f'(x'_p) \\ \widetilde{p}^2_{31} &= \alpha_3 f'(x'_p);\ \widetilde{p}^2_{13} = \alpha_1 f'(x'_p) \\ \widetilde{p}^3_{12} &= \alpha_1 f'(x'_p);\ \widetilde{p}^3_{21} = \alpha_2 f'(x'_p) \end{aligned}$$

The second-stage strategies associate an x_i announcement with each set of prices generated by the first stage. The $\widetilde{x}_i$ corresponding to the $\widetilde{p}$'s will be $\widetilde{x}_i = x'_i$. For other prices, the corresponding choices of $\widetilde{x}_i$ are taken as Nash equilibrium strategies of the resulting subgame; a proof that these equilibria exist is in the Appendix.

These strategies form a subgame-perfect equilibrium for our mechanism. We prove this by first showing that these strategies form Nash equilibria in all subgames and then by demonstrating that no agent can gain by changing his or her first-stage strategy.

By definition, the off-equilibrium-path strategies are subgame-perfect. We still need to show that the equilibrium-path strategies (announcing x'_i) form a Nash equilibrium in the subgame.

The choice of $\widetilde{x}_i$'s implies that $\sum(w_i - \widetilde{x}_i) = x'_p$. The equations determining the y good consumption levels are

$$\begin{aligned} \widetilde{y}_1 &= \alpha_1 \cdot f(x'_p) - (\widetilde{p}^2_{31} + \widetilde{p}^3_{21}) \cdot x'_1 + \widetilde{p}^3_{12} \cdot x'_2 + \widetilde{p}^2_{13} \cdot x'_3 \\ \widetilde{y}_2 &= \alpha_2 \cdot f(x'_p) - (\widetilde{p}^1_{32} + \widetilde{p}^3_{12}) \cdot x'_2 + \widetilde{p}^3_{21} \cdot x'_1 + \widetilde{p}^1_{23} \cdot x'_3 \\ \widetilde{y}_3 &= \alpha_3 \cdot f(x'_p) - (\widetilde{p}^1_{23} + \widetilde{p}^2_{13}) \cdot x'_3 + \widetilde{p}^2_{31} \cdot x'_1 + \widetilde{p}^1_{32} \cdot x'_2 \end{aligned} \tag{6}$$

By substituting for the $\widetilde{p}$'s, we get

$$\widetilde{y}_1 = \alpha_1 \cdot f(x'_p) - [\alpha_3 \cdot f'(x'_p) + \alpha_2 \cdot f'(x'_p)] \cdot x'_1 + \alpha_1 \cdot f'(x'_p) \cdot x'_2 + \alpha_1 \cdot f'(x'_p) \cdot x'_3$$

Since $\alpha_1 + \alpha_2 + \alpha_3 = 1$ and $f'(x'_p) = 1/q$, we can rewrite the equation as

$$\widetilde{y}_1 = \alpha_1 \cdot f(x'_p) + \frac{1}{q}\left[\alpha_1 \cdot \left(\sum x'_i\right) - x'_1\right]$$

Notice that since $\sum x'_i = \sum w_i - x'_p$,

$$x'_1 + q \cdot \widetilde{y}_1 = \alpha_1 \cdot q \cdot f(x'_p) - \alpha_1 \cdot x'_p + \alpha_1 \cdot \left(\sum w_i\right) \tag{7}$$

From Equations (5) and (7) and the fact that $\overline{w}_1 = \alpha_1 \cdot (\sum w_i)$, we see that $\widetilde{y}_1 = y'_1$. We can show that $\widetilde{y}_2 = y'_2$ and $\widetilde{y}_3 = y'_3$ in a similar manner. Note that the outcome is interior and no punishment is invoked.

To show that the $\widetilde{x}_i$ strategies form a Nash equilibrium in the subgame, we must show that no agent has incentive to deviate. Since agent 1 would not like to move the game to a punishment stage, we must show that agent 1 is choosing a strategy that would solve the following:

$$\underset{x_1}{\text{Max}}\, u_1 \left[x_1, \alpha_1 \cdot f(x'_p) - (\widetilde{p}^2_{31} + \widetilde{p}^3_{21}) \cdot x_1 + \widetilde{p}^3_{12} \cdot x'_2 + \widetilde{p}^2_{13} \cdot x'_3\right]$$

The first-order condition for this problem is $u_{11} = u_{12}(\alpha_1 \cdot f'(x'_p) + \widetilde{p}^2_{31} + \widetilde{p}^3_{21})$, which reduces to $u_{11} = u_{12} f'(x'_p)$; the choice of $x_1 = x'_1$ indeed satisfies it. Similarly, it can be shown that announcing x'_2 and x'_3 solves the problems faced by agents 2 and 3. Hence the three-tuple of strategies (x'_1, x'_2, x'_3) constitutes a Nash equilibrium for the second stage of the game.

Finally, we now need to show that no agent would want to deviate in the first stage. To show this for agent 1, first note that agent 1 cannot affect the prices $\widetilde{p}$ he or she faces in the second stage. Second, notice that a change in his or her announced prices would at most change the $\widetilde{x}_2$ and $\widetilde{x}_3$ chosen by the other agents. If the change induces a punishment phase, agent 1 is certainly worse off. However, even if no punishment phase is induced, agent 1 is not better off. Since $\widetilde{p}^3_{12} = \alpha_1 f'(x'_p)$ and $\widetilde{p}^2_{13} = \alpha_1 f'(x'_p)$, the choices of $x_2 = x'_2$ and $x_3 = x'_3$ are optimal for him or her. Similarly, it can be shown no other agent would like to deviate from his or her choice of prices in the first stage, and hence the suggested strategies form a subgame-perfect equilibrium. ■

5. CONCLUSION

We have solved a standard surplus-sharing problem by constructing a feasible two-stage mechanism whose subgame-perfect equilibria exist and give rise to efficient outcomes. While the full description of the mechanism may seem complex, the equilibrium path played does not involve the punishment phases. If the parties know that punishment should never be reached, their analysis of the game becomes substantially simpler.

One should note that the equilibrium allocations in our mechanism are independent of the initial distribution of endowments. They do depend on the α_i parameters, which can be interpreted as ownership shares of the firm in our auxiliary economy. The choice of these parameters can be guided by either notions of fairness (by setting $\alpha_i = w_i/W$, the ownership shares are proportional to initial endowments) or notions of equity (by setting $\alpha_i = 1/3$). A proportional division is often used in firm creation where each partner expects to be rewarded according to the amount of capital he or she invests. An equity-based division is employed in a kibbutz, where the members contribute their initial wealth when joining and share equally in the output. Our mechanism is flexible enough to handle both scenarios as well as intermediate cases by leading to efficient outcomes.

Extending this mechanism to work for more than three agents and two goods is straightforward. It also may be possible to apply our construct that guarantees feasibility and works with general preferences to other externality problems such as those discussed in Varian (1994). The mechanism does not work for environments with increasing returns. This is due to the fact that agents choose prices in the first stage to influence the other agents' second-stage quantity choices. With increasing returns, these desired quantities are not interior, and existence of equilibria is no longer guaranteed. Addressing this issue would require major changes and is a topic of further research. Finally, one also may want to consider environments with different information structures that may lead to games with asymmetrically informed agents.

APPENDIX

We want to show that given any set of prices announced in stage 1, the second-stage game will always have a Nash equilibrium. Rather than analyze our game directly, we will analyze a different game. We will show that an equilibrium always exists in this new game and then show that by adding further strategies that do not upset the equilibrium, we can reach our original game.

In the new game, the x_i outcome coincides with the announced x_i, and except for two contingencies, the following equations determine the y_i outcomes of stage two:

$$\begin{aligned} y_1 &= \alpha_1 \cdot f(x_p) - (p_{31}^2 + p_{21}^3) \cdot x_1 + p_{12}^3 \cdot x_2 + p_{13}^2 \cdot x_3 \\ y_2 &= \alpha_2 \cdot f(x_p) - (p_{32}^1 + p_{12}^3) \cdot x_2 + p_{21}^3 \cdot x_1 + p_{23}^1 \cdot x_3 \\ y_3 &= \alpha_3 \cdot f(x_p) - (p_{23}^1 + p_{13}^2) \cdot x_3 + p_{31}^2 \cdot x_1 + p_{32}^1 \cdot x_2 \end{aligned} \tag{8}$$

The two contingencies are (1) if the sum of the x_i's is greater than W, x_p is zero, and (2) if the RHS of any equation in (8) is less than zero, the corresponding y_i is zero. Note that this game does not imply that the feasibility condition of $\sum y_i \leq f(x_p)$ holds. Since preferences are convex and differentiable and the production function is concave and differentiable, each agent's payoff is continuous and quasi-concave in his or her decision variable. The strategy spaces are convex. If we furthermore impose that the x_i announced cannot exceed some large positive integer $(W + 1)$, the bounded game has an equilibrium. Using standard arguments, it can be shown that this is an equilibrium even in the unbounded game, and therefore, the new game has an equilibrium.

In this new game, agents can guarantee a strictly positive consumption in both goods by announcing a small enough x_i. If $x_p = 0$ and an agent cannot rectify it by changing his or her x_i announcement, it must be the case that another agent j has announced a positive x_j. Since prices are strictly positive and an x_j is positive, a small enough x_i would suffice to generate strictly positive consumption in both goods. In the case where x_p is strictly positive or can be made strictly positive by reducing x_i, the claim is clear.

An x_p equaling zero would imply that one of the y_i's also must be zero (the y_i's are nonnegative and sum to zero). Since an agent can guarantee a strictly positive consumption in both goods, it must be the case that in an equilibrium $x_p > 0$.

Notice that the equilibrium of this game is feasible. We can then remove strategy choices that lead to infeasible outcomes without affecting the equilibrium. This leaves us with an abstract game (see Debreu, 1982). We can then add the punishment phases to return to our original game without destroying this equilibrium, since no agent would have an incentive to enter the punishment phases. One would never want to cause the first punishment phase to be invoked because by doing so one would receive a boundary allocation. Also, one would never want the second punishment phase to punish oneself, and one would never want to cause it to be invoked and punish another agent, since doing so shrinks one's budget set.

REFERENCES

DEBREU, G., "Existence of Competitive Equilibrium," in K. Arrow, and M. Intriligator (eds.), *Handbook of Mathematical Economics*, Vol. 2 (Amsterdam: North Holland, 1982), 697–743.

DUTTA, B., A. SEN, AND R. VOHRA, "Nash Implementation Through Elementary Mechanisms in Economic Environments," *Economic Design* 1 (1995), 173–203.

HURWICZ, L., "Optimality and Informational Efficiency in Resource Allocation Processes," in K. Arrow, S. Karlin, and P. Suppes (eds.), *Mathematical Methods in the Social Sciences* (Stanford, CA: Stanford University Press, 1960), 27–46.

———, "On Informationally Decentralized Systems," in C. B. McGuire, and R. Radner (eds.), *Decision and Organization* (Amsterdam: North Holland, 1972), 297–336.

KAPLAN, T., AND D. WETTSTEIN, "Cost Sharing: Efficiency and Implementation," Foerder Institute for Economic Research working paper no. 44-96, Tel Aviv University, 1996.

MOORE, J., "Implementation, Contracts, and Renegotiation in Environments with Complete Information," in J.-J. Laffont (ed.), *Advances in Economic Theory*, Vol. 1 (Cambridge, England: Cambridge University Press, 1992), 182–282.

MOORE, J., AND R. REPULLO, "Subgame Perfect Implementation," *Econometrica* 56 (1988), 1191–220.

MOULIN, H., "Joint Ownership of a Convex Technology: Comparison of Three Solutions," *Review of Economics Studies* 57 (1990), 439–52.

——— AND S. SHENKER, "Serial Cost Sharing," *Econometrica* 60 (1992), 1009–37.

——— AND A. WATTS, "Two Versions of the Tragedy of the Commons," *Economic Design* 2 (1997), 399–421.

ROEMER, J., AND J. SILVESTRE, "The Proportional Solution for Economies with both Private and Public Ownership," *Journal of Economic Theory* 59 (1993), 426–44.

SAIJO, T., Y. TATAMITANI, AND T. YAMATO, "Toward Natural Implementation," *International Economic Review* 37 (1996), 949–79.

SUH, S.-C., "A Mechanism Implementing the Proportional Solution," *Economic Design* 1 (1995), 310–17.

TIAN, G., "Implementation of Linear Cost Share Equilibrium Allocations," *Journal of Economic Theory* 64 (1994), 568–84.

VARIAN, H., "A Solution to the Problem of Externalities When Agents Are Well Informed," *American Economic Review* 84 (1994), 1278–93.

WALKER, M., "A Simple Incentive Compatible Scheme for attaining Lindahl Allocations," *Econometrica* 49 (1981), 65–71.

YOUNG, H. P., "Cost Allocation," in R. J. Aumann, and S. Hart (eds.), *Handbook of Game Theory*, Vol. II (Amsterdam: North-Holland, 1994), 1193–1235.