Open access • Journal Article • DOI:10.2307/3003619

Incentive Structures Maximizing Residual Gain Under Incomplete Information

Leonid Hurwicz, Leonard D. Shapiro

Published on: 01 Jan 1978 - The Bell Journal of Economics (Center for Economic Research, Department of Economics, University of Minnesota)

Topics: Regret

INCENTIVE STRUCTURES MAXIMIZING RESIDUAL GAIN UNDER INCOMPLETE INFORMATION

by

Leonid Hurwicz *
and
Leonard Shapiro **

Discussion Paper No. 77-83, April 1977

---

* Supported by NSF Grant GS31276X

** Supported by NSF Grant MCS76-05250

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minnesota 55455

ABSTRACT

Two agents are involved in our model. The first agent is to announce a schedule of rewards (or, equivalently, charges) which is a function of the amount produced by the second agent. Then the second agent will decide, using utility maximization, how much to produce. Knowing only the form of the second agent's utility and production functions--not the exact values of their parameters--the first agent seeks to choose a schedule which maximizes the minimum (over all possible utility and productivity parameter values) of a quantity related to his residual gain (residual gain being that part of output remaining after rewards have been paid out). We show that in a broad class of cases the only such maximum is a schedule which takes one-half of production. It should be noted that this result is valid even when schedules are allowed to have certain kinks and/or discontinuities, so that such discontinuities and kinks do not yield any special incentive properties in our model.

This problem is motivated by situations in which the first agent may be thought of as the government and the residual gain (revenue from taxation) is to be used for a paramount national or social objective, e.g., defense to ensure national survival; in this case the second agent represents the country's labor force to be rewarded so as to stimulate a degree of effort maximizing the residual available for national defense. Another possible interpretation is with first agent as a landlord, the second as sharecropper, with value added as the "product" and the problem, seen from the landlord's point of view, being that of maximizing his share of value added.

INTRODUCTION

Our analysis covers at least two models, one from the private and one from the public sector. We first discuss the former: a landlord owns land which can produce a single good, amounts of which are denoted y. The land is worked by a sharecropper. The landlord must choose and announce a schedule of rental fees (charges) for the use of his property. The schedule takes the form of a "reward" function $\rho(y)$ prescribing that if y units of the good are produced[†] from the land, then $\rho(y)$ units are to be kept by the sharecropper, and $y - \rho(y)$ units (the "residual gain") paid to the landlord.

The sharecropper's state of satisfaction depends on two factors: the level of effort expended in producing the good and the amount of reward received.[††] The reward is determined by the output y through the reward function ρ. We shall also make assumptions which imply that effort expended is a (single-valued) function of output y. Thus, indirectly, the sharecropper's state of satisfaction is determined by the level of output y. We denote by $u(y)$ the level of satisfaction obtained by producing y units of output, and call u the ("indirect") utility function of the sharecropper.

It is postulated that the sharecropper will produce b units where b is a point at which u assumes its global maximum. We call b the worker's optimal output. (Thus "optimal" does not mean "Pareto optimal.")[†††]

[†] In a more realistic interpretation, y is the total value added, resulting from the operation of the land by the sharecropper.

[††] It is assumed that the landlord is unable to observe the level of the worker's effort. It is for this reason that the reward is postulated to be a function of output as the variable the landlord can observe.

[†††] See Remark 1.6.

It will be convenient to assume until the end of Section 1 (but not in any proofs) that u has a unique global maximum b, so the optimal output b is uniquely determined. Since the utility of producing y units will depend on the reward $r = \rho(y)$ received by the sharecropper for producing y units, as well as on the effort he must expend, the utility u(y) depends on ρ. On the other hand, the landlord's choice of ρ depends on u because the landlord will use his knowledge about the sharecropper to select ρ so as to maximize his profit.

In previous work,[†] one of us has investigated the extent to which these opposed interests (in our model they are the interests of the utility-maximizing worker and of the residual gain-maximizing landlord) result in a determinate outcome. In that work it was shown that, among linear (fixed share) reward schemes, a 50-50 split is best for the landlord under certain assumptions concerning production and utility functions. Also, a conjecture concerning rewards other than linear ones was stated. Here we confirm that conjecture and obtain further results. After making some restrictive assumptions concerning production and utility functions (less restrictive than those in the earlier work) we will prove that if the landlord is motivated solely by his own interests and knows nothing about the sharecropper's utility function, and he has at his disposal a class of reward functions ("piecewise smooth" functions) which is much broader than the linear reward functions,then his best action is still to share with the sharecropper, 50-50, the proceeds of production.

Another model encompassed by our analysis is that of a community threatened by an outside danger. The community's total output of goods

---

[†]An early version was presented by Hurwicz at the Conference on the Economics of Internal Organization at the University of Pennsylvania, September 19-21, 1974. See also Hurwicz [1977].

and services y (which we shall treat as if it were one-dimensional) must be divided, between consumption and defense of the community's existence. This is done by announcing a reward function $\rho(y)$ to be applied to individual output,[†] i.e., an individual worker who produces y units will retain $\rho(y)$ for his own consumption and $y - \rho(y)$, the "residual gain," will go toward the community's defense. As in the previous model, each worker is assumed to maximize a utility u which is indirectly determined by the level of output. Here the community, like the landlord of the previous example, wants to pick ρ so as to maximize its residual gain, but the choice of such a ρ depends on u. And the worker's utility, maximized with respect to y, depends on ρ. This example differs from the first in two important respects: first, it is natural to assume that the workers, whose paramount desire is that the community survive, will want the community to maximize the residual gain available to cope with the danger,[††] whereas the sharecropper is presumably not interested in the landlord's maximizing his residual gain. Second, the lack of information on the part of the community about which u's are to be faced is due to the variety of workers as well as to their desire for privacy.

For simplicity we will use, in the remainder of this paper, the terms landlord and worker.

The reader interested in other formulations of incentive problems in sharecropping may refer to works by Bardhan and Srinivasan (1971), Bell and Zusman (1976), Cheung (1969), and Stiglitz (1974). Approaches to incentive problems which are related to our model can be found in works by Keren (1969), Leibenstein (1966), Marschak (1976), and Mirlees (1973). Models with analogous structure have been studied by Moiseev (1975) and Vatel and Erezhkov (1973).

[†] Again, value added might be a more appropriate interpretation of y.

[††] The "paradox" is that, despite this desire, the worker's individual effort is assumed to be at a level maximizing his own utility rather than the community's welfare. But concern for the community's survival would presumably make the worker vote in favor of adopting a reward function ρ maximizing the residual gain.

1. DEFINITIONS, A SPECIAL CASE, AND AN OUTLINE OF THE PAPER

1.1 Definitions: utilities U and u, disutilities $\varphi \in \Phi$, optimal output b and profit π.

We let z denote the worker's effort and r the reward which the worker receives.[†] We assume the worker's utility function can be written in the form $U(r,z) = r - \psi(z)$.[*] We assume the production function $y = f(z)$ to be invertible,[††] so setting $\psi \circ f^{-1} = \varphi$ and recalling that the reward r is given by $\rho(y)$, we have

$$u(y) = U(\rho(y), f^{-1}(y)) = \rho(y) - \varphi(y).$$

It is to be noted that the "indirect" utility function u is determined by the "direct" utility function U, the production function f, and the reward function ρ.

We call the term $\varphi(y)$ appearing in the representation $u(y) = \rho(y) - \varphi(y)$ the disutility term and call φ the disutility for short. We assume that the landlord knows nothing about which disutility φ appears in the worker's utility function except that φ is a member of a certain set Φ. Many of our definitions and results will depend on what we choose Φ to be, i.e., on what we assume the landlord knows about the worker.

If the production function f were of the constant returns to scale type, $f(z) = cz$ for some $c > 0$, and the disutility term ψ (in the direct utility function U) were quadratic, $\psi(z) = dz^2$ for some $d > 0$, then we would have $\varphi(y) = \psi \circ f^{-1}(y) = \alpha y^2$ where $\alpha = d/c^2$. We denote by φ_α the function $\varphi_\alpha(y) = \alpha y^2$ for $y \geq 0$. We refer to the disutilities φ_α for $\alpha > 0$

[†] r, y, and utilities are real numbers throughout the paper. z is a real number in our examples, but much of our analysis is applicable to a multi-dimensional z (see next footnote).

[††] In fact, we need only assume that for each $y \geq 0$, the set $\{\psi(z): f(z) = y\}$ has a mininum. With this assumption it is more plausible to view z as ranging over a multidimensional or other space, and our results are applicable provided that the function $\varphi \circ F$ has the assumed properties.

[*] See Proposition 7.4 for the case $U(r,z) = r^\gamma - \psi(z)$, $\gamma \leq 1$.

as quadratic disutilities.

Given a reward function ρ and disutility φ we define $b(\rho,\varphi)$ to be the global maximizer (assumed unique[†]) with respect to $y \geq 0$ of $u(y) = \rho(y) - \varphi(y)$, and we call $b(\rho,\varphi)$ the worker's optimal output. Since ρ will be understood from the context we will suppress it and write $b(\varphi)$, the optimal output of the worker characterized by φ. Given ρ and φ we define the landlord's residual gain (or, for short, gain), to be $\pi(\rho,\varphi) = b(\varphi) - \rho(b(\varphi))$. We may write π for $\pi(\rho,\varphi)$. One would expect the landlord to choose a reward schedule ρ so as to maximize his gain π. In general, "maximizing π" is an inadequate criterion for choosing (see 1.11). but we next discuss a special case where it is adequate.

1.2 The special case of linear rewards

Let us consider the special case where the landlord must choose his reward function ρ from among the linear reward functions, i.e., those of the form $\rho_k(y) = ky$ for some $k \geq 0$, and where the set Φ of disutilities is that of the quadratics, $\Phi = \{\varphi_\alpha : \alpha > 0\}$.

1.3 Lemma. For any real numbers $k \geq 0$ and $\alpha > 0$ we have

$$b(\rho_k,\varphi_\alpha) = \frac{k}{2\alpha} \quad \text{and} \quad \pi(\rho_k,\varphi_\alpha) = \frac{1}{2\alpha}(k - k^2).$$

Proof. The optimal output $b(\rho_k,\varphi_\alpha)$ is defined to be the global maximizer of $u(y) = ky - \alpha y^2$ with respect to $y \geq 0$ and that quadratic function assumes its maximum at $\frac{k}{2\alpha}$. The gain $\pi(\rho_k,\varphi_\alpha)$ is defined to be

---

[†]This assumption is made mainly for simplicity of exposition and is not needed in any proofs (see Remark 2.3). We drop this uniqueness assumption in Sections 3, 4, and 5, where our main results are proved.

$$b(\rho_k,\varphi_\alpha) - \rho_k(b(\rho_k,\varphi_\alpha)) = \frac{k}{2\alpha} - k(\frac{k}{2\alpha}) = \frac{1}{2\alpha}k(1 - k).$$

The next lemma is obvious but we will need to refer to it often.

1.4 Lemma. The maximum of $k(1 - k)$, with respect to k, occurs only at $k = \frac{1}{2}$. Thus the inequality $\frac{1}{4} \leq k(1 - k)$ has only one solution, namely, $k = \frac{1}{2}$.

The next result follows from the previous two lemmas. It was first proved in [Hurwicz, 1977].

1.5 Theorem. For each $\alpha > 0$ the maximum of $\pi(\rho_k,\varphi_\alpha)$, with respect to $k \geq 0$, occurs only at $k = \frac{1}{2}$. Thus if the landlord is restricted to choosing linear reward functions ρ_k, $k \geq 0$, and faces quadratic disutilities φ_α, $\alpha > 0$, he can maximize gain only by choosing the reward function $\rho^*(y) = \frac{y}{2}$ for all $y \geq 0$.

We will use the symbol ρ^* throughout the paper for the reward function defined by $\rho^*(y) = \frac{y}{2}$ for $y \geq 0$.

1.6 Remark. Because this paper deals with the objective of maximizing the residual gain of the landlord it is important to note that, in general, the "joint welfare" of the landlord and the worker is not being maximized. In particular, the solution just offered ($k = \frac{1}{2}$) does not yield an allocation that is Pareto optimal for the worker and the landlord. It is not difficult to see that, among the linear solutions $\{\rho_k : k \geq 0\}$, the only one that is Pareto optimal gives to the worker all of the output ($k = 1$).

1.7 Outline of the paper

The previous theorem says that if the landlord (1) must choose his reward schedule from the linear ones ρ_k, (2) faces only quadratic disutilities φ_α, and (3) seeks a reward schedule ρ which maximizes $\pi(\rho,\varphi_\alpha)$ for each α, then his only choice is ρ^*. It is the aim of this paper to generalize the theorem with respect to the hypotheses (1) and (2), but this will require our weakening the criterion (3) of maximizing π. We will still obtain the conclusion that ρ^* is the only choice satisfying the weakened criterion.

We denote by P the set of reward functions from which the landlord may choose. Thus, in Theorem 1.5, $P = \{\rho_k: k \geq 0\}$, the linear rewards.

In the remainder of Section 1, we will describe the largest class P of rewards to which our results apply. We will call the members of this largest class "permissible"; they are essentially the piecewise C^1 functions. We will also consider various criteria which the landlord might apply in choosing a reward function, and will settle on what we call "efficiency," which requires that ρ maximize the infimum over Φ of efficiency ratios (not gains). The efficiency ratio is defined to be the ratio of gain to the supremum of possible profits assuming complete information, and is analogous to Savage's "regret." That ρ^* is an efficient reward will imply that it is undominated. Most of the remainder of the paper is devoted to proving that ρ^* is the unique efficient reward function in a very general setting. In order to motivate and clarify the lengthy proof that ρ^* is the unique efficient reward in this general setting, we present in Section 2 a special case where the proof is relatively simple. The most crucial assumptions we make in this special case are that Φ is the quadratic disutilities and P consists of rewards satisfying: If $\bar{y} > 0$, then there is a disutility φ_α (i.e., a worker) such that $\bar{y}$ maximizes $u(y) = \rho(y) - \varphi_\alpha(y)$ (i.e., $\bar{y}$ is the worker's optimal output), and $\rho(y)$ is a C^1 function in $y \geq 0$.

It is worth noting that just the assumption that ρ is C^1 would not yield a simple proof; the proof in that case is almost as long as the proof in the general piecewise C^1 ("permissible") case.

In the general case we must discard our assumption on rewards $\rho \in P$ that optimal output $b(\varphi)$ is unique (i.e., that for each $\varphi \in \Phi$, $\rho(y) - \varphi(y)$ has a unique maximizer with respect to $y \geq 0$), and this requires some redefinitions which we give in Section 3. These redefinitions all reduce to the former ones in case $b(\varphi)$ is unique for all $\varphi \in \Phi$. In Sections 4 and 5 we prove that ρ^* is the unique efficient reward when P is the largest set considered, the permissible rewards, and Φ is the quadratic disutilities. This is done by proving the result, in Section 4, for an artificially contrived special case, then, in Section 5, reducing the general case to the one in Section 4.

All the cases discussed in Sections 1-5 assume Φ is the quadratic disutilities, $\Phi = \{\varphi_\alpha : \alpha > 0\}$. Section 6 is devoted to seeing how much we can weaken this assumption on Φ. It is easily proved (6.1) that ρ^* remains undominated as long as Φ contains $\{\varphi_\alpha : \alpha > 0\}$ as a subset. But ρ^* will not remain the unique efficient reward function if we enlarge Φ arbitrarily. We show this by an example (6.3). We then show (6.7) that ρ^* is the unique efficient reward if $\Phi \supseteq \{\varphi_\alpha : \alpha > 0\}$ and each φ in Φ satisfies the following conditions: $\varphi'(0) = 0$ and φ, φ', φ'' positive on the interval $(0,\infty)$ and $\varphi'' \geq 0$ on $(0,\infty)$ (primes denote derivatives).

In section 7 we study four cases, in each of which ρ^* is dominated by another permissible reward function. The first two cases result from bounding (above or below) the values of α which may appear in the worker's disutility term φ_α . In the third case we replace $\varphi_\alpha(y) = \alpha y^2$ by the function $\varphi_\alpha^\beta(y) = \alpha y^\beta$ for some $\beta > 1$, and in the fourth case we replace the direct utility $U(r,z) = r - \psi(z)$ by $U_\gamma(r,z) = r^\gamma - \psi(z)$ for some γ , $0 \le \gamma \le 1$. Although ρ^* is dominated in each case, it turns out that, in the first three cases, if the landlord has somewhat less information available then the results of sections 5 and 6 apply and ρ^* is the unique (among linear rewards) efficient reward. In the fourth case the linear reward given by $\rho(y) = y/4$ for $y \ge 0$ is the unique (among linear rewards) efficient reward function.

1.8 Which rewards ρ can the landlord choose?

Empirically, we often observe (possibly with y as value added) that P is either the linear functions $\{\rho_k : k > 0\}$ (in sharecropping, for example) or the piecewise affine functions (in income tax schedules, for example).† We will be more tolerant and define the most general class, of "permissible" reward functions, to be the piecewise C^1 functions, i.e., those having a continuous derivative except on a closed discrete set--see Figure 1(b).

Before defining the permissible reward functions we review some standard notation. For any real function f, we write $f(x^+) = \lim_{t \to x^+} f(t)$, and $f(x^-) = \lim_{t \to x^-} f(t)$. Also, $f'_+(x) = \lim_{t \to x^+} \frac{f(t) - f(x)}{t - x}$, and similarly for $f'_-(x)$. Thus $f'_+(x)$, the right derivative of x, may not equal $f'(x^+)$, the right limit of ordinary derivatives of x. Clearly. f is continuous at x iff $f(x^-)$ and $f(x^+)$ exist and equal f(x). We say f has a jump discontinuity at x if $f(x^-)$ and $f(x^+)$ both exist and are unequal. A function f of a real variable is said to be nonincreasing if $x < y$ implies $f(x) \geq f(y)$, nondecreasing if $x < y$ implies $f(x) \leq f(y)$. $\mathbb{R}_+$ denotes the set of nonnegative reals. We say a function f defined on $\mathbb{R}_+$ is C^1 at zero if f'_+ exists at 0 and equals $f'(0^+)$.

With this notation and terminology we can define "permissible" rigorously:

---

†To explain our terminology "piecewise affine," we note that an affine function is one of the form $\rho(y) = ky + \ell$ for constants k and ℓ; an affine function is linear when $\ell = 0$. We say a function ρ is piecewise affine if it is affine on each one of a countable set of intervals and these intervals cover the real line except for a closed discrete set of points. A closed discrete set is one having a finite number of points in any finite interval. Thus the set of integers is closed discrete while the set $\{\frac{1}{n}: n = 1,2,\ldots\}$ is not, whether or not zero is added to it. (See Figure 1(a), where points in the discrete closed set are marked a, b, c.)

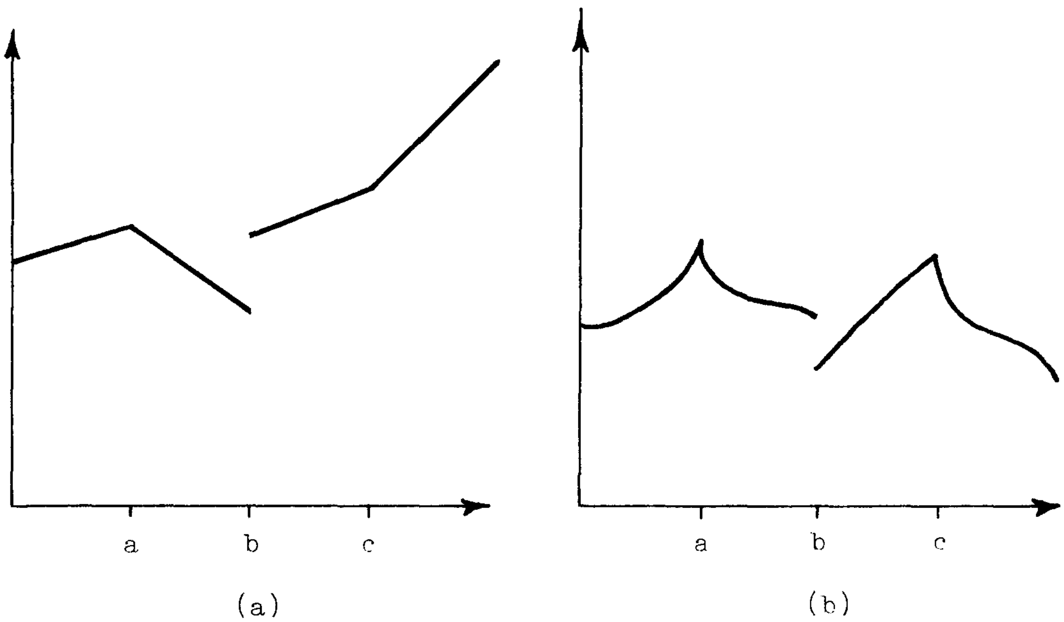


Figure 1

1.9 Definition. The function $\rho(y)$: $\mathbb{R}_+ \to \mathbb{R}_+$ is permissible if and only if:

(a) $0 \leq \rho(y) \leq y$ for all $y \geq 0$.

(b) $\rho(y)$ is piecewise C^1, i.e., it is continuously differentiable except on a closed discrete set X. Furthermore, $0 \notin X$.

(c) The discontinuities (if any) of ρ are jump discontinuities.

(d) If ρ is discontinuous at $y \in X$, then $\rho(y) = \max\{\rho(y^+), \rho(y^-)\}$.

Part (a) of this definition is a natural assumption to make about reward functions which the landlord can choose, namely, that he will never reward the worker with more than is produced and that the rewards are always nonnegative. With a more careful analysis we might be able to omit the condition $\rho(y) \leq y$: if we consider a reward function $\bar{\rho}$ such that $\bar{\rho}(y_0) > y_0$ for some y_0, it is reasonable to expect that the landlord could find another permissible reward function $\bar{\bar{\rho}}$ with $\bar{\rho}(y_0) > \bar{\bar{\rho}}(y_0) > y_0$, such that each worker produced the same amount under $\bar{\bar{\rho}}$ as $\bar{\rho}$, yet $\bar{\bar{\rho}}$ would cost the landlord potentially less in rewards. Concerning the hypothesis $\rho(y) \geq 0$, if we allowed $\rho(y)$ to be negative, then the landlord could "enslave" workers, for example with the schedule $\rho(y) = -1$ for all y, in which case the worker would always have to pay the landlord all of his output plus one unit of the good.

Part (c) of this definition and the assumption $0 \notin X$ have been added to simplify the analysis--they are probably not essential.

Part (d) of the definition means that at a discontinuity the worker gets the more advantageous of the two limiting values $\rho(y^+)$, $\rho(y^-)$. As is proved in the following proposition, part (d) of the definition implies that if disutilities are quadratic, then u(y) has a maximum for any permissible ρ, i.e., that every worker, indexed by some α, has an optimal output $b(\varphi_\alpha)$(= maximizer of u). It is easy to show that without part (d) such a maximum might not exist, in which case it would not be clear how to define optimum, profit, etc.

1.10 Proposition (Existence of optimizing outputs). If ρ is permissible and the disutilities are quadratic, then for every $\alpha > 0$ there is at least one point $b(\varphi_\alpha) \geq 0$ such that $u(b(\varphi_\alpha)) = \sup \{u(y): y \geq 0\}$, i.e., such that u attains its global maximum at $b(\varphi_\alpha)$.

Proof. Since $\rho(y) \leq y$ by part (a) of Definition 1.9, we know that

$$u(y) = \rho(y) - \varphi_\alpha(y) \leq y - \alpha y^2 = y(1 - \alpha y).$$

This implies that $u(y) \leq 0$ for $y \geq \frac{1}{\alpha}$. Since part (a) with $y = 0$ implies $u(0) = 0$, we conclude that it suffices to prove the existence of some $b(\varphi_\alpha) \in [0, \frac{1}{\alpha}]$ such that $u(b(\varphi_\alpha)) = \sup \{u(y): y \in [0, \frac{1}{\alpha}]\}$. But part (d) implies that ρ is upper semicontinuous, so u is upper semicontinuous, and any upper semicontinuous function on a compact interval such as $[0, \frac{1}{\alpha}]$ attains its maximum. Q.E.D.

1.11 Criteria for choosing a reward function

It will be helpful to view our model as a game, where ρ (the landlord's strategy) must be chosen from P, φ (the worker's strategy) must be chosen from Φ, $\pi(\rho,\varphi)$ is the payoff to the landlord, and $u(b(\rho,\varphi))$ the worker's payoff. In the language of game theory, Theorem 1.5 says ρ^* is a dominant strategy for the landlord when $P = \{\rho_k: k \geq 0\}$ and $\Phi = \{\varphi_\alpha: \alpha > 0\}$.

1.12 Definition. Given sets P and Φ and a (payoff) function $\pi: P \times \Phi \to \mathbb{R}$, and $\rho, \rho^0 \in P$, we say ρ^0 dominates ρ over Φ for the function π if $\pi(\rho^0,\varphi) \geq \pi(\rho,\varphi)$ for all $\varphi \in \Phi$ and strict inequality holds for some $\varphi \in \Phi$. We say ρ^0 is dominant (with respect to Φ, P and π) if for every $\rho \in P$, ρ^0 dominates ρ over Φ for π, and we say ρ^0 is undominated $(w,r,t,\Phi,P$ and $\pi)$ if no $\rho \in P$ dominates ρ^0 over Φ.

In the case of linear rewards ρ_k and quadratic disutilities φ_α we **(1.5)** found a dominant reward function ρ^* which "maximized profit," but in the more general cases we will consider there will not exist a dominant reward function for the payoff function π. To see the problem more clearly, refer to Figure 2 where we have graphed $\pi(\rho,\varphi_\alpha)$ as a function of $\frac{1}{8\alpha}$ for different functions ρ. Recall from (1.3) that $\pi(\rho_k,\varphi_\alpha) = \frac{1}{2\alpha}(k - k^2)$, and $k - k^2$ attains its maximum at $k = \frac{1}{2}$, so the graph of $\pi(\rho^*,\varphi_\alpha)$ dominates (is above) that of $\pi(\rho_k,\varphi_\alpha)$ for $k \neq \frac{1}{2}$, for all $\alpha > 0$. But other rewards ρ lead to graphs of $\pi(\rho,\varphi_\alpha)$ such as those marked ρ^0, ρ^{00} in the graph. In each of these cases the reward function dominates ρ^* for extreme values of $\frac{1}{8\alpha}$ (specifically, ρ^0 dominates ρ^* when $\frac{1}{8\alpha}$ is near 0 and ρ^{00} dominates ρ^* when $\frac{1}{8\alpha}$ tends to $+\infty$) but each of ρ^0 and ρ^{00} is dominated by ρ for other values of $\frac{1}{8\alpha}$.

Although we cannot expect ρ^* to be dominant in general, we should at least expect it to be undominated, i.e., there should not exist $\rho \in P$ such that $\pi(\rho,\varphi_\alpha) \geq \pi(\rho^*,\varphi_\alpha)$ for all α. This follows from our results.

One criterion we could ask ρ^* to meet[†] in a more general setting is that it maximize the infimum of gains

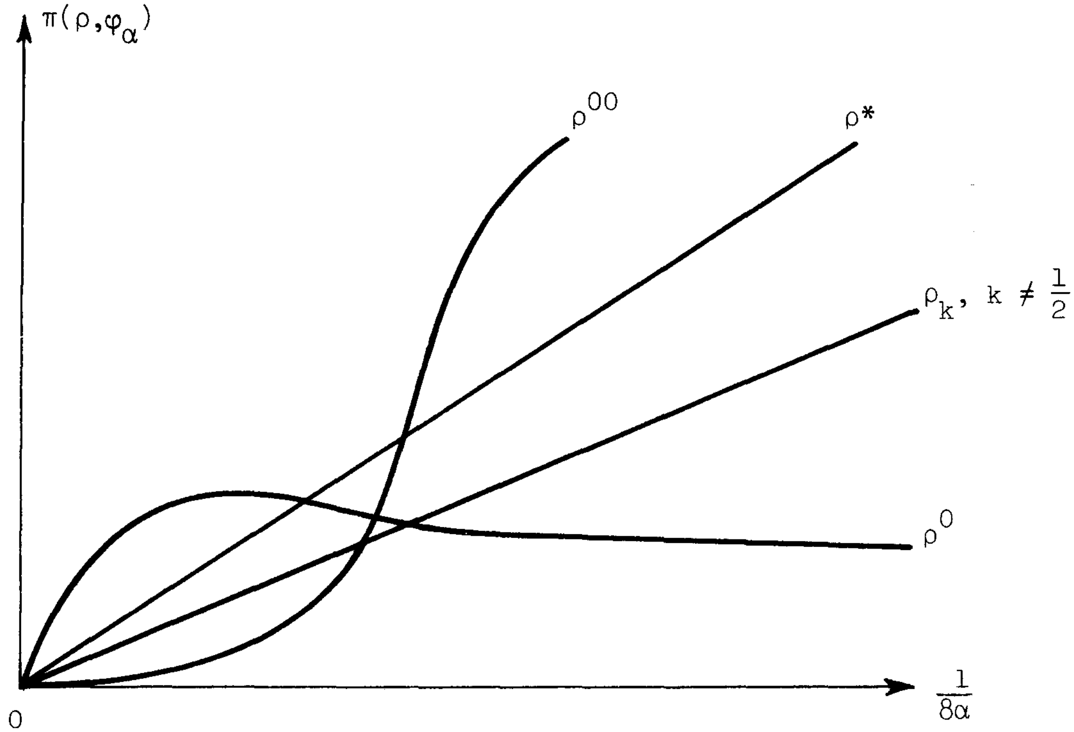
$$(1.13) \qquad \inf \{\pi(\rho,\ ): \varphi \in \Phi\}.$$

Unfortunately, this will yield no worthwhile result, since the infimum in (1.13) is always zero for the ρ's and Φ's we will consider.[††] This problem is not unfamiliar in game theory and statistics, and one response to it (Savage, 1954, p.163) has been to use "regret" instead of "payoff." The regret is usually computed by finding the best payoff assuming complete

---

[†]Another criterion, maximization of expected profit with respect to a probability distribution which is assumed given, has been used by others such as Stiglitz and Mirlees.

[††]This is because as $\alpha \to \infty$, the maximum of $u(y) = \rho(y) - \alpha y^2$ occurs at points $b(\varphi_\alpha)$ converging to zero and then $\pi = b(\varphi_\alpha) - \rho(b(\varphi_\alpha))$ converges to zero.



Graphs of $\pi(\rho,\varphi_\alpha)$ as a function of $\frac{1}{8\alpha}$, for various ρ's.

Figure 2

information and subtracting the payoff from it. We use the regret approach in this paper. However, we will use division instead of subtraction because of the multiplicative relationships between quantities in our model. In this form, regret can be interpreted as a measure of efficiency of a policy. We acknowledge that the regret principle does have certain disadvantages (see, for example, Chernoff, 1954).

In order to compute our version of regret, we define the "best payoff assuming complete information" by:

$$\hat{\pi}_P(\varphi) = \sup \{\pi(\rho,\varphi): \rho \in P\} \text{ for } \varphi \in \Phi.$$

It follows from Theorem 1.5 that in the case $P = \{\rho_k: k \geq 0\}$, $\hat{\pi}_P(\varphi_\alpha) = \frac{1}{8\alpha}$ for $\alpha > 0$. In all other cases we consider, $\hat{\pi}_P(\varphi_\alpha) = \frac{1}{4\alpha}$ for all $\alpha > 0$ (see 2.2). Thus a landlord who knows φ can do twice as well (in fact, appropriate virtually the "total surplus"--see the proof of Lemma 2.2) when not confined to linear reward functions. On the other hand, it will be seen below that in terms of our efficiency (regret) criterion the landlord derives no advantage from being permitted to use nonlinear reward functions when he is ignorant of φ. Notice that $\pi/\hat{\pi}$ measures the "efficiency" of the reward system.

1.14 Definition. A reward function ρ^0 is efficient with respect to P and Φ if $\rho^0 \in P$ and ρ^0 maximizes

$$\inf \{\pi(\rho,\varphi)/\hat{\pi}_P(\varphi): \varphi \in \Phi\}$$

among all $\rho \in P$.

As mentioned above, we will show that in several cases ρ^* is the unique efficient reward in P. In turn, this uniqueness implies that ρ^* is undominated;

in fact it implies that ρ^* is undominated for both payoff functions π and $\pi/\hat{\pi}$. To see this for $\pi/\hat{\pi}$, suppose on the contrary that ρ dominates ρ^*, i.e.,

(1.15) $\quad \pi(\rho,\varphi)/\hat{\pi}(\varphi) \geq \pi(\rho^*,\varphi)/\hat{\pi}(\varphi)$ for all $\varphi \in \Phi$.

Then, certainly

$$\inf\ \{\pi(\rho,\varphi)/\hat{\pi}(\varphi):\ \varphi \in \Phi\} \geq \inf\ \{\pi(\rho^*,\varphi)/\hat{\pi}(\varphi):\ \varphi \in \Phi\}$$

so ρ^* could not be the unique efficient reward. To see that ρ^* is undominated for π, note that the truth of (1.15) remains unchanged if we multiply through by $\hat{\pi}(\varphi)$, which is positive for all φ we consider.

2. A SOMEWHAT SPECIAL CASE

Although it is our goal to prove our claim that ρ^* (defined by $\rho^*(y) = y/2$ for $y \geq 0$) is the unique efficient **reward, in as general a setting as possible, we feel that it will be helpful to present here a somewhat special (but broader than linear) case. The proof in this special case is** rather simple, but the general proof (see Section 4) parallels it quite closely, thus we feel that an understanding of this special case will greatly facilitate following the rather lengthy proof in the general case. None of the results from this section except Lemma 2.2 are used in subsequent sections.

2.1 Theorem

Suppose the set Φ *of disutilities is that of the quadratics,*

$$\Phi = \{\varphi_\alpha : \alpha > 0\}$$

where $\varphi_\alpha(y) = \alpha y^2$ *for* $y \geq 0$, *and* P *is the set of reward functions* ρ *satisfying*

2.1.1 $\rho(y)$ is a C^1 function for $y \geq 0$

2.1.2 $\rho(0) = 0$

2.1.3 Given an amount of production $\bar{y} > 0$, there is an $\alpha > 0$ such that $\bar{y}$ maximizes $u(y) = \rho(y) - \varphi_\alpha(y)$ with respect to $y \geq 0$.

2.1.4 For every $\alpha > 0$ there is a unique maximizer $b(\varphi_\alpha)$ of $u(y) = \rho(y) - \varphi_\alpha(y)$ with respect to $y \geq 0$, and $b(\varphi_\alpha) > 0$ for $\alpha > 0$.

Then ρ^* *is the unique efficient reward function with respect to* P *and* Φ.

Before proving 2.1 we will compute $\hat{\pi}_P(\varphi_\alpha)$. This computation will be used again in a later section.

2.2 Lemma. For any set P of rewards such that $\rho(0) \geq 0$ for all $\rho \in P$, $\hat{\pi}_P(\varphi_\alpha) \leq \frac{1}{4\alpha}$ for all $\alpha > 0$. If P is defined by 2.1.1-2.1.4, then $\hat{\pi}_P(\varphi_\alpha) = \frac{1}{4\alpha}$ for all $\alpha > 0$.

Proof. First, we claim that if $\rho(0) \geq 0$ and if $b = b(\varphi_\alpha)$ is an optimal output, i.e., b maximizes $u(y) = \rho(y) - \varphi_\alpha(y)$ with respect to $y \geq 0$, then the landlord's gain is no greater than $b - \alpha b^2$ (a term we might call the "total surplus"). This is because

$$b - \alpha b^2 = [b - \rho(b)] + [\rho(b) - \alpha b^2]$$

and the first bracketed term is the landlord's gain while the second is the worker's utility. Utility at b must be nonnegative since $u(b) \geq u(0)$ by the assumption of utility maximization and $u(0) = \rho(0) - \alpha 0^2 \geq 0$. Nonnegativity of utility implies that $b - \alpha b^2 \geq b - \rho(b)$, as claimed.

Simple calculus shows that the maximum of $b - \alpha b^2$ occurs at $b = 1/2\alpha$ and is $1/4\alpha$. Thus we conclude that if each $\rho \in P$ satisfies $\rho(0) \geq 0$, then $\hat{\pi}_P(\varphi_\alpha) \leq \frac{1}{4\alpha}$.

Now suppose P is defined by 2.1.1-2.1.4. Then we claim that given $\alpha > 0$ we can, by an appropriate choice of $\rho \in P$, get the worker with disutility φ_α to produce near $\frac{1}{2\alpha}$ and the gain to be arbitrarily close to $\frac{1}{2\alpha} - \alpha(\frac{1}{2\alpha})^2 = \frac{1}{4\alpha}$. This will prove $\hat{\pi}_P(\varphi_\alpha) = \frac{1}{4\alpha}$. The φ we choose is illustrated in Figure 3 below: ρ itself in 3(a) and its derivative ρ' in 3(b). The function ρ is, for some $\varepsilon > 0$, equal to $\alpha y^2 + \varepsilon y$ on the interval $[0, \frac{1-\varepsilon}{2\alpha}]$ and to $y - \frac{(1-\varepsilon)^2}{4\alpha}$ on the interval $[\frac{1-\varepsilon}{2\alpha}, \infty)$.

Recalling that

$$\rho(y) - \alpha y^2 = \int_0^y (\rho'(t) - 2\alpha t)\, dt,$$

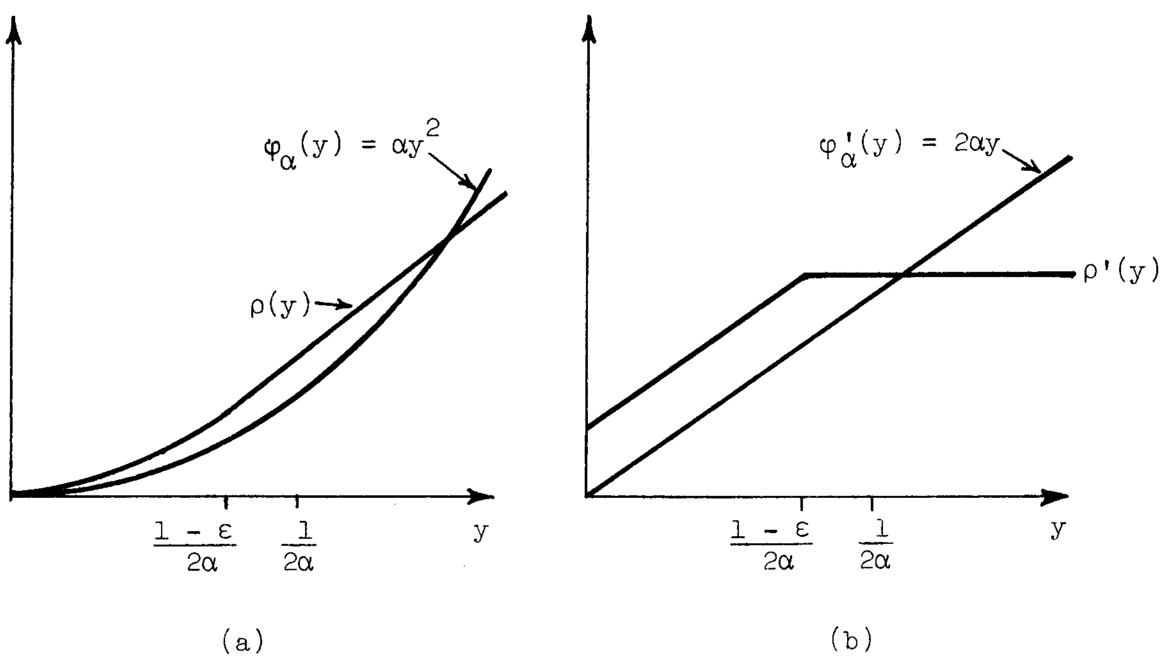

$\varphi_\alpha(y) = \alpha y^2$
$\rho(y)$
$\frac{1-\varepsilon}{2\alpha}$
$\frac{1}{2\alpha}$
y
(a)
$\varphi'_\alpha(y) = 2\alpha y$
$\rho'(y)$
$\frac{1-\varepsilon}{2\alpha}$
$\frac{1}{2\alpha}$
y
(b)

Figure 3

one can easily see from Figure 3(b) that the maximum of $\rho(y) - \alpha y^2$ occurs at $b = \frac{1}{2\alpha}$. The gain for this b is

$$b - \rho(b) = b - (b - \frac{(1-\varepsilon)^2}{4\alpha}) = \frac{(1-\varepsilon)^2}{4\alpha},$$

and as $\varepsilon \to 0$ this converges to $\frac{1}{4\alpha}$. All that remains is to show that the ρ defined above and pictured in Figure 3 satisfies conditions 2.1.1-2.1.4. **Conditions 2.1.1-2.1.2 are clear, and using graphs like those in Figure 3 one easily checks that conditions 2.1.3-2.1.4 are satisfied (but note that the fixed α which is used to define ρ here is different than the variable α appearing in the conditions 2.1.3,4).**

Proof of Theorem 2.1. For ρ^* to be the unique efficient reward means, in this case, that it is the only reward which maximizes $\inf\{\pi(\rho,\varphi_\alpha)/\hat{\pi}(\varphi_\alpha): \alpha > 0\}$. Thus we must prove that if $\rho \in P$ and ρ satisfies

(0) $\inf\{\pi(\rho,\varphi_\alpha)/\hat{\pi}(\varphi_\alpha): \alpha > 0\} \geq \inf\{\pi(\rho^*,\varphi_\alpha)/\hat{\pi}(\varphi_\alpha): \alpha > 0\}$,

then $\rho = \rho^*$. By 1. and the previous lemma, (0) is equivalent to:

$$\inf\{4\alpha\pi(\rho,\varphi_\alpha): \alpha > 0\} \geq \frac{1}{2}.$$

Since $\pi(\rho,\varphi_\alpha) = b(\varphi_\alpha) - \rho(b(\varphi_\alpha))$, this is in turn equivalent to:

(1) $4\alpha[b(\varphi_\alpha) - \rho(b(\varphi_\alpha))] \geq \frac{1}{2}$ for all $\alpha > 0$.

Thus the proof will be completed if we can show that $\rho \in P$ and (1) imply $\rho = \rho^*$.

Since $b(\varphi_\alpha)$ maximizes $\rho(y) - \alpha y^2$ with respect to $y \geq 0$ and ρ is C^1, the first-order conditions hold:

$$\rho'(b(\alpha)) = 2\alpha b(\alpha) \qquad \text{for all } \alpha > 0.$$

(The equality holds because $b(\alpha) > 0$ by assumption 2.1.4, so $b(\alpha)$ is not a corner solution.) We substitute $2\alpha = \rho'(b(\alpha))/b(\alpha)$ in (1) to get

$$\frac{1}{4} \leq \rho'(b(\alpha))[1 - \frac{\rho(b(\alpha))}{b(\alpha)}] \qquad \text{for all } \alpha > 0.$$

By hypothesis 2.1.3 each $y > 0$ is equal to $b(\varphi_\alpha)$ for some $\alpha > 0$, so we conclude

(2) $$\frac{1}{4} \leq \rho'(y)[1 - \frac{\rho(y)}{y}] \qquad \text{for all } y > 0.$$

Take the limit as $y \to 0$ in (2), and use our assumption that $\rho(0) = 0$, to get:

$$\frac{1}{4} \leq \rho'(0)[1 - \rho'(0)],$$

so, by 1.4 ,

(3) $$\rho'(0) = \frac{1}{2}, \quad \text{i.e.,} \frac{\rho(y)}{y} \to \frac{1}{2} \text{ as } y \downarrow 0.$$

Since ρ' is continuous, and never zero by (2), (3) implies

(4) $$\rho'(y) > 0 \qquad \text{for } y \geq 0.$$

Next, we claim

(5) $$\rho'(y) \geq \frac{\rho(y)}{y} \qquad \text{for } y \geq 0.$$

The inequality (5) says marginal reward is never below average reward. It would follow from standard arguments (as would the rest of this proof) if we had assumed convexity of ρ. To prove (5) we suppose, seeking a contradiction, that for some $y > 0$, $\rho'(y) < \rho(y)/y$. Apply this to (2) to get

$$\frac{1}{4} \leq \rho'(y)[1 - \frac{\rho(y)}{y}] < \rho'(y)[1 - \rho'(y)]$$

which is impossible by Lemma 1.4. Next, we claim that

(6) $\frac{\rho(y)}{y}$ is nondecreasing.

This is because, by (5),

$$\frac{d}{dy}\left(\frac{\rho(y)}{y}\right) = \frac{1}{y}\left(\rho'(y) - \frac{\rho(y)}{y}\right) \geq 0.$$

By (2) and (4), $\rho(y)/y$ is bounded above by 1, so for some finite L,

(7) $\frac{\rho(y)}{y} \uparrow L \qquad$ as $y \to \infty$.

We claim that there do not exist y_0 and $\delta > 0$ such that

(8) $\rho'(y) > \frac{\rho(y)}{y} + \delta \qquad$ for $y > y_0$.

If (8) held, then by (7) we could pick y_0 large enough so that

(9) $\rho'(y) > L + \frac{\delta}{2} \qquad$ for $y > y_0$.

But this is impossible for it would imply

$$L = \lim_{y\to\infty} \frac{\rho(y)}{y} = \lim_{y\to\infty} \frac{1}{y}\int_0^{y_0} \rho'(t)\, dt + \lim_{y\to\infty} \frac{1}{y}\int_{y_0}^{y} \rho'(t)\, dt$$

$$\geq 0 + \lim_{y\to\infty} \frac{1}{y}\int_{y_0}^{y} \left(L + \frac{\delta}{2}\right) dt = L + \frac{\delta}{2}.$$

Thus there is no $\delta > 0$ satisfying (8). This and (5) imply that for some sequence $\{y_n\} \to \infty$,

(10) $\lim_n \rho'(y_n) = \lim_n \frac{\rho(y_n)}{y_n}$.

Apply this to (2) and use (7):

$$\frac{1}{4} \leq \lim_n \rho'(y_n)[1 - \frac{\rho(y_n)}{y_n}] = L[1 - L].$$

By Lemma 1.4 , $L = \frac{1}{2}$. Now (3), (6), and (7) imply $\rho(y)/y \equiv \frac{1}{2}$, as desired.

<u>2.3 Remark</u>. Hypothesis 2.1.3 is quite strong--it is what simplifies the proof of 2.1. See Figure 4 in Section 3 for an example where it does not hold, even though ρ is C^1. The existence and uniqueness of optimal output $b(\varphi_\alpha)$--2.1.4--would follow if we strengthened 2.1.2 to $0 \leq \rho(y) \leq y$ for all y: existence is proved above in 1.10 and uniqueness could be derived from Lemma 4.1 below.

3. NONUNIQUENESS OF OPTIMAL OUTPUT $b(\varphi)$ AND THE REDEFINITIONS IT NECESSITATES

Until Section 6 we will discard our assumption that $b(\varphi)$ is unique, i.e., that for each $\rho \in P$ and $\varphi \in \Phi$, $\rho(y) - \varphi(y)$ has a unique maximizer with respect to $y \geq 0$. This will require a more complicated notation. It also happens that until Section 6 we will only consider cases in which the set of disutilities Φ is the quadratics, and this will allow us to simplify our notation somewhat since we will be able to use $b(\alpha)$ in place of $b(\varphi_\alpha)$. In this section we will discuss the effects of these changes.

Notice that in the two special cases we have considered thus far, optimal output $b(\varphi_\alpha)$ has been unique: in 1.2 , the linear case, this was because the utility $u(y) = ky - \alpha y^2$ was quadratic so had a unique maximizer, and in 2.1 it followed from our assumption (2.1.3) that all possible positive amounts of production were covered--see the remark at the end of Section 2.

Proposition 1.10 shows that if ρ is permissible, then for each $\alpha > 0$ the utility function $u(y) = \rho(y) - \varphi_\alpha(y)$ has at least one maximizer, but what if it has more than one? In that case the worker would have some leeway in deciding how much optimally to produce. To explain this situation it will help to introduce some notation. Fix a permissible ρ, and for $\alpha > 0$ let $B(\alpha)$ denote the set of global maximizers, with respect to y, of $u(y) = \rho(y) - \alpha y^2$. By definition, B is a correspondence and since there may be more than one such maximizer for a given α, it is not necessarily a function. A listing of choices by each worker, of how much to produce, is a selection $b(\cdot)$ from the correspondence $B(\cdot)$, and we call such a function $b(\cdot)$ an "optimal selection" which is <u>compatible</u> with ρ (see 3.1 below).

To see visually what is happening, consider the reward function in Figure 4(a), whose derivative is graphed in Figure 4(b). The correspondence

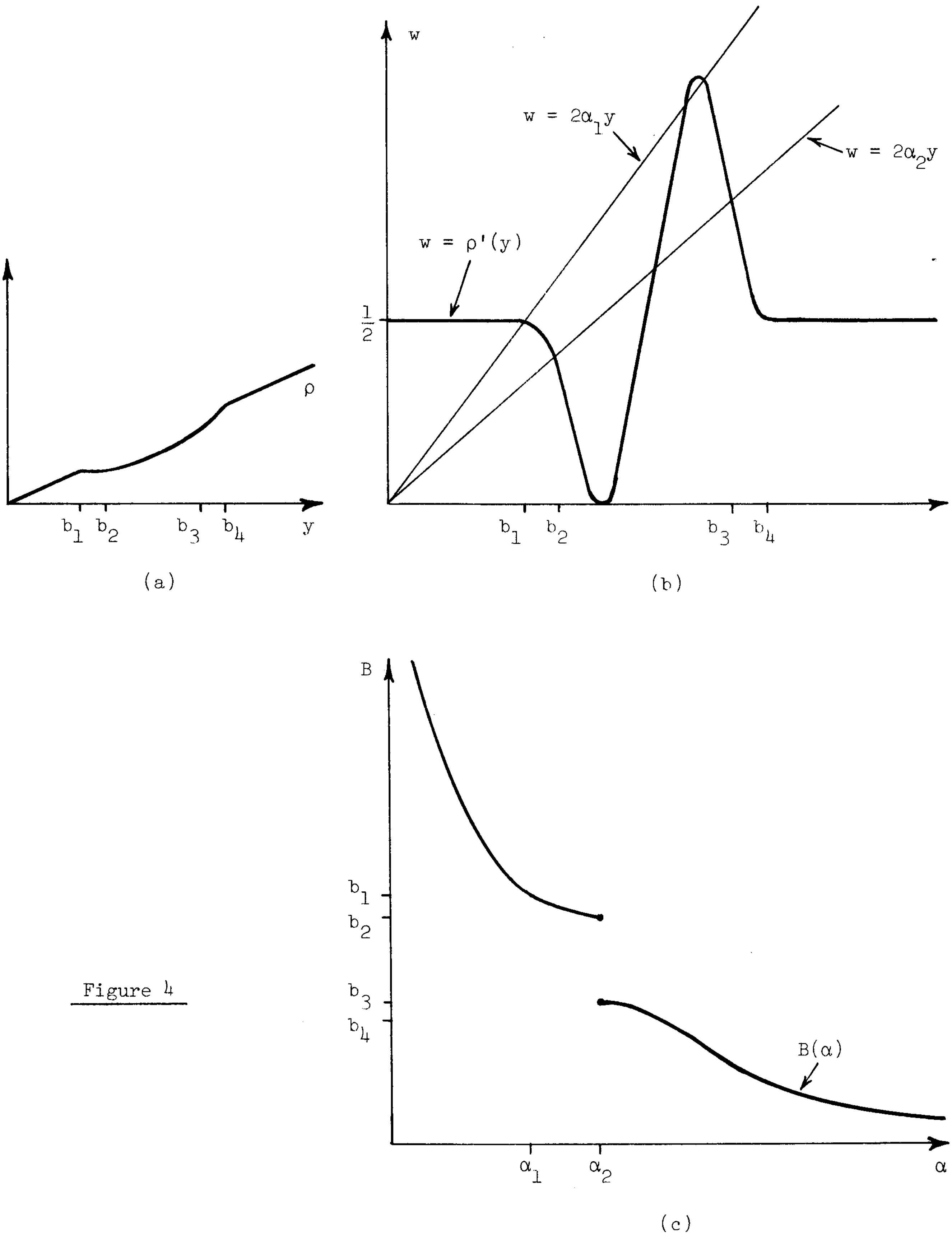

ρ
b1 b2 b3 b4 y
(a)
w
w = 2α1y
w = 2α2y
w = ρ'(y)
1/2
b1 b2 b3 b4
(b)
B
b1
b2
b3
b4
B(α)
α1 α2 α
(c)

Figure 4

$B(\cdot)$ related to ρ is graphed in Figure 4(c). The reward function $\rho(y)$ is meant to be equal to $y/2$ on the intervals $(0,b_1)$ and (b_4,∞) and to dip as indicated on (b_1,b_4). Thus, by 1.3 , $B(\alpha) = \{\frac{1}{8\alpha}\}$ on the intervals $(0,b_1)$ and (b_4,∞), since the dip in (b_1,b_4) is not large enough to change the behavior outside (b_1,b_4). To understand these graphs it is helpful to keep in mind the first-order condition which says that if b maximizes the C^1 function $\rho(y) - \alpha y^2$, then $\rho'(b) = 2\alpha b$, i.e., the graphs of $w = \rho'(y)$ and $w = 2\alpha y$ intersect over the point $y = b$. Also, keep in mind that since the ρ in the figure is C^1,

$$\rho(y) - \alpha y^2 = \int_0^y (\rho'(t) - 2\alpha t)\, dt,$$

so maximizing $\rho(y) - \alpha y^2$ is equivalent to maximizing the area between the graphs of $w = \rho'(y)$ and $w = 2\alpha y$. Notice that for $\alpha = \alpha_1$, as marked in 4(b) and 4(c), there is a unique maximizer, marked b_1, but for $\alpha = \alpha_2$, there are two maximizers, namely, b_2 and b_3.

When we made assumptions implying that $u(y)$ had a unique maximizer for each α, $B(\cdot)$ was a function so there was only one optimal selection $b(\cdot)$ compatible with ρ. In the more general case the gain $\pi = b(\alpha) - \rho(b(\alpha))$, which was previously a function of ρ and α, will be a function of ρ and $b(\cdot)$ and α. That is to say the landlord's gain will depend not only on the α-value of the worker faced but also on which one of the worker's utility-maximizing b's he decides to produce. This alters the definition of the efficiency ratio $\pi/\hat{\pi}$ and of an efficient reward function. We now give these broader definitions, to be used until Section 6, all of which reduce to the former definitions in case disutilities are quadratic and $u(y) = \rho(y) - \alpha y^2$ has a unique maximizer for each α. Since we will be dealing only with permissible reward functions until Section 6 (see 1.9 for the definition of permissible) we will not explicitly refer to the set P of rewards.

<u>3.1 Definition</u>. The function $b(\cdot)$: $\mathbb{R}_+ \to \mathbb{R}_+$ is <u>compatible</u> with the reward function ρ if, for every $\alpha > 0$, $b(\alpha)$ is a global maximizer for $u(y) = \rho(y) - \alpha y^2$ with respect to y, i.e.,

$$(3.2) \qquad \rho(b(\alpha)) - \alpha[b(\alpha)]^2 \geq \rho(y) - \alpha y^2 \qquad \text{for all } y \geq 0.$$

Given a permissible reward function ρ and $b(\cdot)$ compatible with ρ, we define the landlord's <u>residual gain</u> or, briefly, <u>gain</u>, from a worker indexed by α, to be

$$\pi = \pi(\rho,b(\cdot),\alpha) = b(\alpha) - \rho(b(\alpha));$$

and for $\alpha > 0$ we define $\hat{\pi}$ by

$$\hat{\pi}(\alpha) = \sup\{\pi(\rho,b(\cdot),\alpha): \rho \text{ is permissible, } b(\cdot) \text{ is compatible with } \rho, \text{ and } \alpha > 0\}.$$

We define the <u>efficiency ratio</u> of a permissible reward function ρ at $b(\cdot)$ and $\alpha > 0$ to be:

$$e(\rho,b(\cdot),\alpha) = \pi(\rho,b(\cdot),\alpha)/\hat{\pi}(\alpha)$$

(notice we have introduced the new notation e), and we say a permissible reward function ρ is <u>efficient</u> if for every $b(\cdot)$ compatible with ρ we have

$$\inf\{e(\rho,b(\cdot),\alpha): \alpha > 0\} \geq \inf\{e(\bar{\rho},\bar{b}(\cdot),\alpha): \alpha > 0\}$$

for every permissible $\bar{\rho}$ and $\bar{b}(\cdot)$ compatible with $\bar{\rho}$. (Of course, this definition of efficient is with respect to the set Φ of quadratic disutilities and the set P of permissible rewards.)

4. FIRST STEP IN THE GENERAL PROOF: REDUCING TO A SIMPLER CASE

In this section we begin the proof that in our most general case, ρ^* (defined by $\rho^*(y) = y/2$ for $y \geq 0$) is the unique efficient reward function. Here and in Section 5 we will prove this theorem for the case where P, the set of reward functions from which the landlord may choose, is that of the permissible functions (defined in 1.9--these are essentially the piecewise C^1 functions) and where Φ, the set of disutility functions which the landlord may face, is that of the quadratics, $\Phi = \{\varphi_\alpha : \alpha > 0\}$ where φ_α is defined by $\varphi_\alpha(y) = \alpha y^2$ for $\alpha \geq 0$. In Section 6 we will extend the result to larger sets Φ.

As one would expect from the special case proved in 2.1, the proof that ρ^* is the unique efficient reward with respect to P and $\Phi = \{\varphi_\alpha : \alpha > 0\}$ easily reduces to showing that if $\rho \in P$ and $b(\cdot)$ is compatible with ρ, then the hypothesis

$$(4.1^*) \qquad \frac{1}{4} \leq 2\alpha[b(\alpha) - \rho(b(\alpha))] \qquad \text{for all } \alpha > 0$$

implies $\rho = \rho^*$, i.e., $\rho(y) = y/2$ for all $y \geq 0$. We shall refer to (4.1*) as the "maximality hypothesis." In this section we will establish this key implication for an artificially constructed set P of reward functions. Then, in Section 5, we will reduce the general case (P = permissible rewards) to the one we have considered here.

We begin with a lemma which will also be used in Section 5. Thus we state and prove it for any permissible reward function.

4.1 Lemma. Suppose ρ is permissible (cf. 1.9), $b(\cdot)$ is an optimal selection compatible with ρ, and the maximality hypothesis is satisfied:

$$(*) \qquad \frac{1}{4} \leq 2\alpha[b(\alpha) - \rho(b(\alpha))] \qquad \text{for all } \alpha > 0.$$

Then (i) $b(\cdot)$ is nonincreasing, (ii) $b(\alpha) > 0$ for all $\alpha > 0$, (iii) $b(\alpha) \to 0$ as $\alpha \to \infty$, (iv) $b(\alpha) \to \infty$ as $\alpha \to 0^+$.
(v) (First Order Condition) If ρ is C^1 at $b(\alpha)$ then $\rho'(b(\alpha)) = 2\alpha b(\alpha)$.

Proof: (i) Let $\alpha < \beta$; we must prove $b(\alpha) \geq b(\beta)$. By compatibility (3.2) applied to $b(\beta)$ and $y = b(\alpha)$, then to $b(\alpha)$ and $y = b(\beta)$, we have

$$\rho(b(\beta)) - \beta\, b(\beta)^2 \geq \rho(b(\alpha)) - \beta\, b(\alpha)^2$$

$$\rho(b(\alpha)) - \alpha\, b(\alpha)^2 \geq \rho(b(\beta)) - \alpha\, b(\beta)^2 \quad .$$

Adding these together and rearranging, we get

$$[\beta\, b(\alpha)^2 - \alpha\, b(\alpha)^2] - [\beta\, b(\beta)^2 - \alpha\, b(\beta)^2] \geq 0 \quad ,$$

which is of the form

$$(*) \quad [\beta\, s^2 - \alpha\, s^2] - [\beta\, t^2 - \alpha\, t^2] \geq 0$$

where $s = b(\alpha)$, $t = b(\beta)$. The expression (*) can be rewritten as

$$(\beta - \alpha)(s^2 - t^2) \geq 0 \quad .$$

Since $\beta > \alpha$, this implies $s \geq t$, or $b(\alpha) \geq b(\beta)$, as desired.

(ii) To prove $b(\alpha) > 0$ for all α , recall our maximality hypothesis:

$$\frac{1}{4} \le 2\alpha[b(\alpha) - \rho(b(\alpha))] \quad .$$

Since ρ is nonnegative by 1.9a, we must have $b(\alpha) > \rho(b(\alpha)) \ge 0$.

(iii) If iii is false then since b is nonincreasing there is an $M > 0$ such that $b(\alpha) \uparrow M$ as $\alpha \to \infty$. By (3.2) for $y = 0$,

$$(**)\ \rho(0) \le \rho(b(\alpha)) - \alpha[b(\alpha)]^2 \quad \text{for all } \alpha \ .$$

Take the limit as $\alpha \to \infty$ in (**) and use 1.9c to get

$$\rho(0) \le \rho(M^+) - \lim_{\alpha\to\infty} \alpha\, b(\alpha)^2 \ .$$

But the Limit is $+\infty$ since $b(\alpha) \to M > 0$, a contradiction.

(iv) We assume, on the contrary, that $b(\alpha) \to M < \infty$ as $\alpha \to 0^+$. By the maximality hypothesis,

$$\frac{1}{4} \le \lim_{\alpha\to 0^+} 2\alpha[b(\alpha) - \rho(b(\alpha))]$$

so we must have

$$+\infty = \lim_{\alpha\to 0} [b(\alpha) - \rho(b(\alpha))] = M - \rho(M^-) \quad ,$$

which is impossible for $M < \infty$.

(v) If ρ is C^1 at $b(\alpha)$ then since $b(\alpha)$ is a maximum with respect to $r \ge 0$ of $\rho(r) - \alpha r^2$, and $b(\alpha)$ is an interior maximum by (ii), the first-order conditions must hold, namely $\rho'(b(\alpha)) - 2\alpha b(\alpha) = 0$.

This completes the proof of the Lemma.

The rest of this section is devoted to proving the following.

4.2 Theorem. Suppose ρ is a reward function, $b(\cdot)$ is compatible with ρ, and they satisfy:

4.2.1 ρ is permissible (defined in 1.9);

4.2.2 ρ is continuous and nondecreasing;

4.2.3 There is a countable set of disjoint intervals $[c_\ell, d_\ell]$ with $\ell = 1,2,\ldots$ such that $c_\ell > 0$ for all ℓ and if $y > 0$ and $y \neq b(\alpha)$ for all $\alpha > 0$, then $y \in [c_\ell, d_\ell]$ for some ℓ. (Some or all of the intervals $[c_\ell, d_\ell]$ may be empty (if $c_\ell > d_\ell$)!)

4.2.4 For every $\ell = 1,2,\ldots,$

$$\rho(y) = \beta_\ell y^2 + k_\ell \qquad \text{for all } y \in [c_\ell, d_\ell]$$

where k_ℓ is a constant and $c_\ell = b(\beta_\ell^+)$, $d_\ell = b(\beta_\ell^-)$.

If ρ and b satisfy the maximality hypothesis

$$(4.1^*) \qquad \frac{1}{4} \leq 2\alpha[b(\alpha) - \rho(b(\alpha))] \qquad \text{for all } \alpha > 0,$$

then $\rho(y) = y/2$ for all $y \geq 0$.

Before beginning the proof of 4.2 we will try to explain and motivate the conditions 4.2.2-4.2.4.

The conditions 4.2.2-4.2.4 are implied by the conditions on ρ given in the "somewhat special case," 2.1. For example, if (by 2.1.3) $b(\cdot)$ is onto $(0,\infty)$, then 4.2.3 and 4.2.4 are satisfied with $[c_\ell, d_\ell]$ empty for all ℓ. It is also possible to show that if $b(\cdot)$ is onto, then ρ is nondecreasing.

The previous comment about 4.2.3 and 4.2.4 shows that they generalize the condition 2.1.3 that $b(\cdot)$ is onto. They do so by imposing a special, quadratic structure on ρ, on the intervals $[c_\ell, d_\ell]$ which $b(\cdot)$ may miss. The purpose of this special structure is best explained by Figure 5.

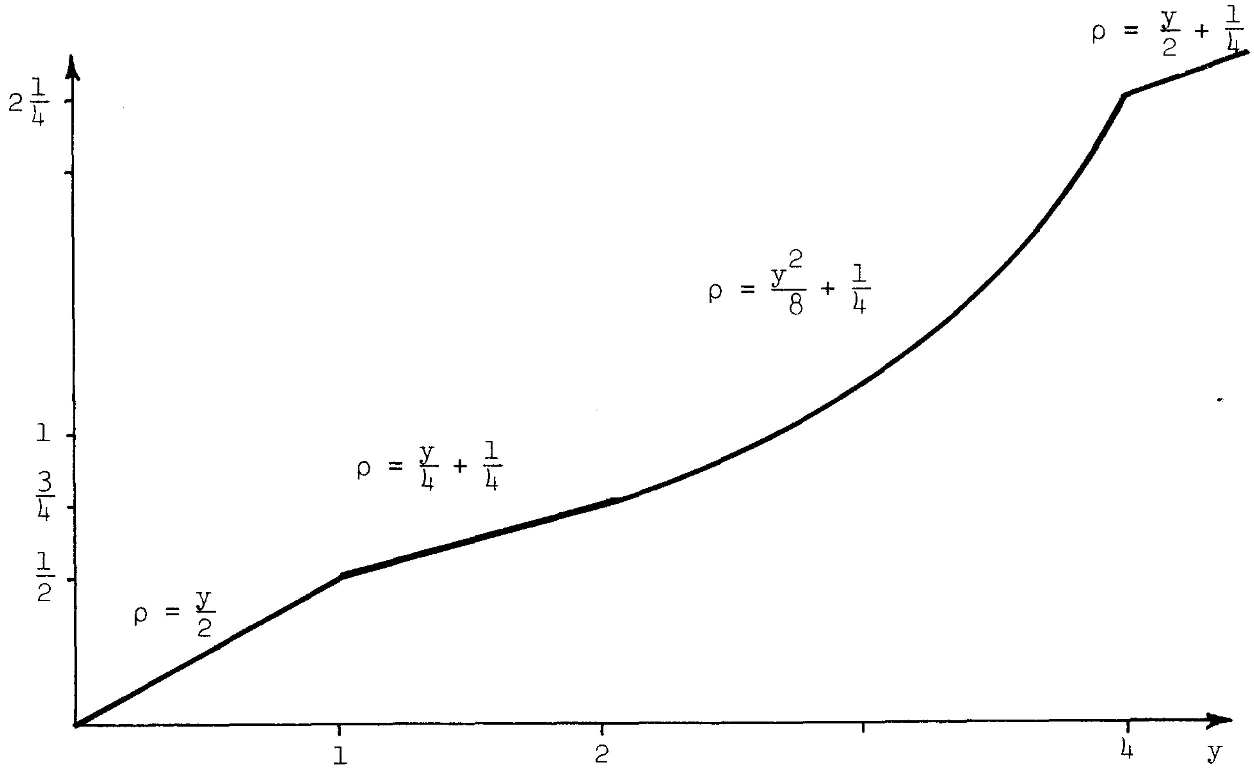


Figure 5(a): Graph of ρ.

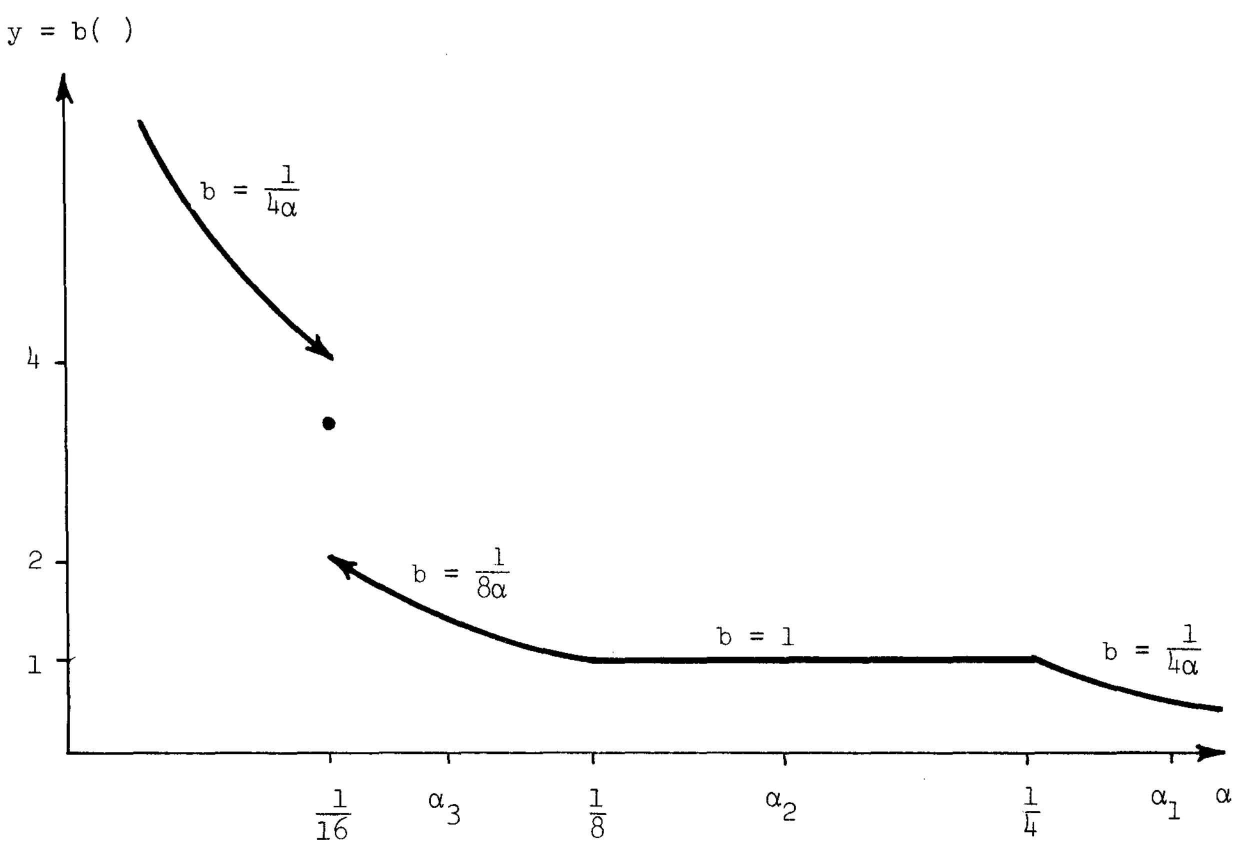


Figure 5(b): Graph of b(•).

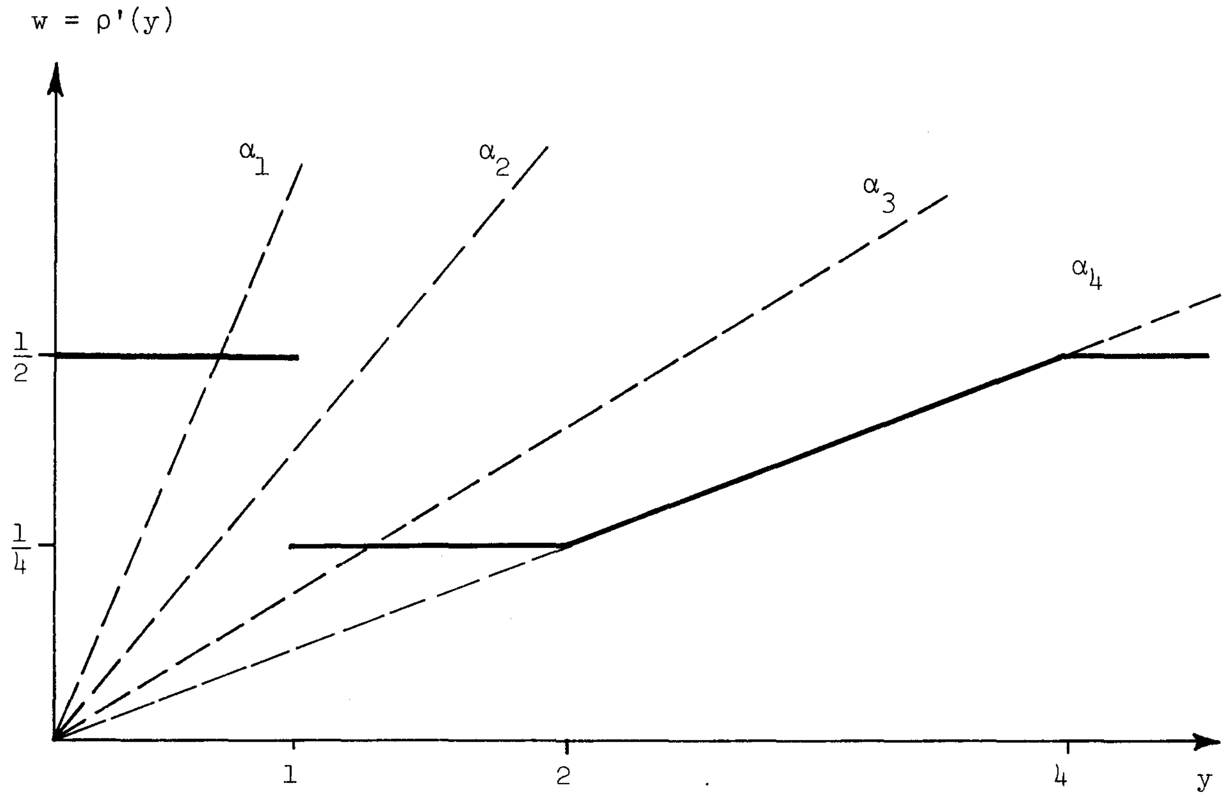


Figure 5(c): Graph of ρ'.

Figure 5a is the graph of a reward function ρ satisfying conditions 4.2, 5b is a selection $b(\cdot)$ compatible with ρ, and 5c is a graph of the derivative ρ' of ρ. It is useful to use lemma 4.1v, the first-order condition, to understand figure 5c, and that figure in turn will help clarify 5a and 5b. The first-order condition 4.1v means that given α, if ρ is C^1 at $b(\alpha)$ then $b(\alpha)$ is a y-value at which the graph of $w = \rho'(y)$ intersects the straight line $w = 2\alpha y$. A few such straight lines are drawn in 5c, labeled α_i, $i = 1,2,3,4$ (i.e. α_i marks the line $w = 2\alpha_i y$) and the corresponding α_i values are marked on 5b.

Notice that the y's which $b(\cdot)$ misses, i.e. those y's such that $y \neq b(\alpha)$ for all $\alpha > 0$, are all in the interval $[2,4]$. The interval $[2,4]$ is where $b(\cdot)$ "jumps" in 5b, at $\alpha = \alpha_4$. As one can see most easily from figure 5c, maxima of $\rho(y) - \alpha_4 y^2$ occur at <u>every</u> point of the interval $[2,4]$. Condition 4.2.4 is designed to ensure that if b "jumps" at β then the value it takes at the jump could be any point on the interval between the limiting values $b(\beta^+)$, $b(\beta^-)$ -- in our example $\beta = \alpha_4$. One other comment may serve to explain 4.2.3 and 4.2.4 further: We shall see in section 5 that the conditions 4.2 are typical in that each permissible ρ is "equivalent" to one satisfying 4.2. This equivalence is analogous to the fact that a consumer's utility function with indifference curves given by the solid line in figure 6a is equivalent to a convex one, as in 6b. By equivalence of the utility functions we mean that if a utility-maximizing choice from a budget set such as the shaded area is consistent with 6a, then it is consistent with 6b.

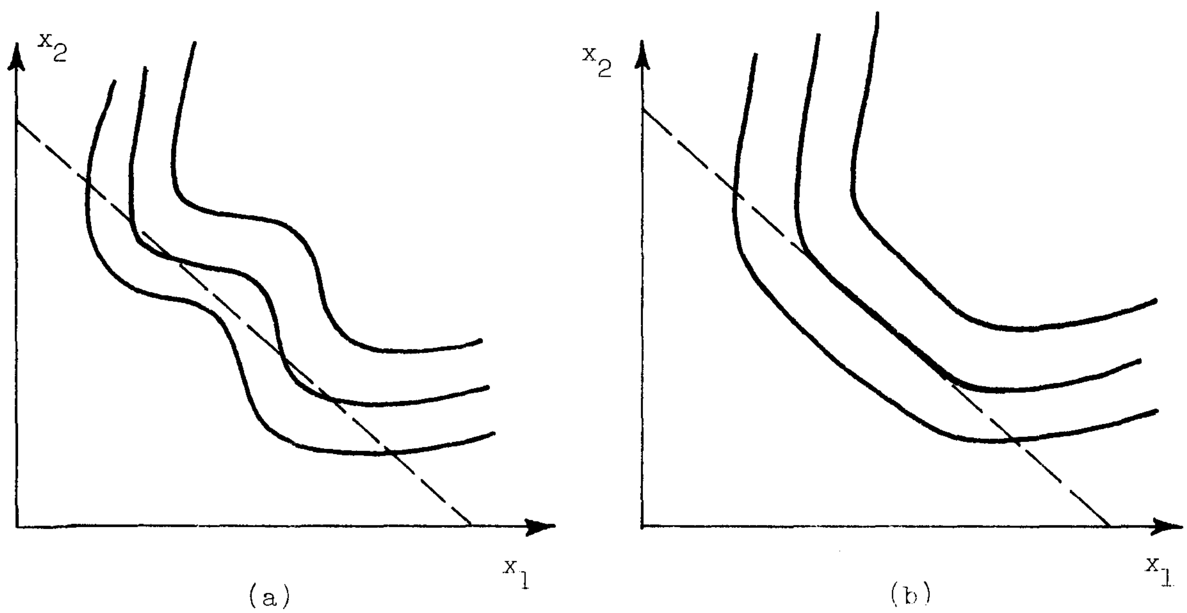


Figure 6: x_1, x_2 = commodities

Now to return to the proof of Theorem 4.2. We assume 4.2.1-4.2.4 and the maximality hypothesis 4.1* and separate the proof into five lemmas:

4.3 Lemma. If ρ is C^1 at $b(\alpha)$, then

$$\frac{1}{4} \leq \rho'(b(\alpha))[1 - \frac{\rho(b(\alpha))}{b(\alpha)}].$$

Proof. By the first-order condition 4.1v, $\rho'(b(\alpha)) = 2\alpha b(\alpha)$. Now the Lemma follows if we substitute $2\alpha = \rho'(b(\alpha))/b(\alpha)$ into the maximizing hypothesis 4.1*, which is permitted since $b(\alpha) > 0$ by 4.1ii.

4.4 Lemma. If ρ is C^1 at y and $y > 0$ then

(i) $\rho'(y) \geq \frac{\rho(y)}{y}$.

Proof: First consider the case when y is in the range of $b(\cdot)$, i.e. $y = b(\alpha)$ for some α . Let us assume, by way of contradition, that

(ii) $\rho'(y) < \rho(y)/y$.

Then we claim

(iii) $\frac{1}{4} \leq \rho'(y)[1 - \rho(y)/y] < \rho'(y)[1 - \rho'(y)]$.

The first inequality follows from 4.3, writing y in place of $b(\alpha)$, the second from (ii) and the fact that $\rho'(y) = 2\alpha b(\alpha) > 0$ by 4.1ii. Now (iii) contradicts (take $r = \rho'(y)$) ,so we have proved the lemma if $y = b(\alpha)$.

We are left with the case y is not in the range of b, i.e., $y \neq b(\alpha)$ for all α, in which case (by 4.2.3 and our assumption $y > 0$), $y \in [c_\ell, d_\ell]$ for some ℓ and $\rho(y) = \beta_\ell y^2 + K_\ell$ for some constant K_ℓ. We denote B_ℓ, K_ℓ, c_ℓ, and d_ℓ by β, K, c, and d for the remainder of the Lemma's proof. If we substitute $2\beta y$ for $\rho'(y)$ and $\beta y^2 + K$ for $\rho(y)$, (i) reduces to

(iv) $\quad \beta y \geq K/y \quad$ for $y \in [c,d]$.

Recall that $\beta > 0$ since $b(\cdot)$ is defined only on positive reals, and $c > 0$ by 4.2.3. Thus if $K \leq 0$, (iv) follows easily. So suppose $K > 0$. Then the left side of (iv), βy, is increasing in y while the right side, K/y, is decreasing. Thus it suffices to prove (iv) for $y = c$. We need to prove

(v) $\quad \beta c \geq K/c$.

Since $c > 0$ by 4.2.3 and ρ is piecewise C^1 by 1.9, there is a sequence $\{\alpha_n\}$ such that ρ is C^1 at $b(\alpha_n)$ for all n and $\alpha_n \downarrow \beta$, $b(\alpha_n) \uparrow c$. Then since $\rho'(b(\alpha_n)) = 2\alpha_n b(\alpha_n)$, we have

$$2\alpha_n b(\alpha_n) = \rho'(b(\alpha_n)) \geq \rho(b(\alpha_n))/b(\alpha_n)$$

from the first paragraph of the proof. Taking the limit on n gives

(vi) $\quad 2\beta c \geq \dfrac{\rho(c)}{c}$,

and substituting $\rho(c) = \beta c^2 + K$ in (vi) gives (v), as desired.

<u>4.5 Lemma</u>. <u>The function</u> $\rho(y)/y$ <u>is nondecreasing in</u> $y > 0$.

<u>Proof</u>. If ρ is C^1 at y, then by 4.4

$$\frac{d}{dy}(\rho(y)/y) = (1/y)(\rho'(y) - \rho(y)/y) \geq 0,$$

so $\rho(y)/y$ is nondecreasing on intervals where it is C^1. But since ρ is continuous and is C^1 except on a discrete closed set (1.9b) this implies $\rho(y)/y$ is everywhere nondecreasing.

<u>4.6 Lemma</u>. $\frac{1}{2} \le \rho(y)/y \le 1$ <u>for all</u> $y > 0$.

<u>Proof</u>: We have assumed (1.9b) that 0 is not in X , the discrete closed set where ρ is not C^1. Thus ρ is C^1 in a one-sided neighborhood of 0 in $\mathbb{R}_+$, and, using the assumption (1.9a with $y=0$) that $\rho(0) = 0$, we have

$$\text{(i)} \quad \lim_{y\to 0^+} \frac{\rho(y)}{y} = \lim_{y\to 0^+} \rho'(y) = \rho'(0) \quad .$$

Since ρ is C^1 in such a one-sided neighborhood and $b(\alpha) \to 0$ as $\alpha \to \infty$ by 4.1iii, for α sufficiently large ρ is C^1 at $b(\alpha)$. Thus by 4.3

$$\text{(ii)} \quad \frac{1}{4} \le \rho'(b(\alpha))[1 - \rho(b(\alpha))/b(\alpha)] \quad \text{for } \alpha \text{ sufficiently large.}$$

We can take the limit as $\alpha\to\infty$ in (ii) and apply (i) to get

$$\frac{1}{4} \le \rho'(0)[1 - \rho'(0)] \quad .$$

Now 1.4 implies $\rho'(0) = \frac{1}{2}$. Since by Lemma 4.5 $\rho(y)/y$ is nondecreasing, this and (i) imply $1/2 \le \rho(y)/y$ for all y .

The other inequality, $\rho(y)/y \leq 1$, is part (1.9a) of the definition of permissibility. But since we feel that this part of the definition can be omitted, we will here prove the inequality independently. If $\rho(y_0)/y_0 > 1$ for some y_0 then by Lemma 4.5 $\rho(y)/y > 1$ for all $y > y_0$. Since by 4.1iv $b(\alpha) \to \infty$ as $\alpha \to 0$ we can find α_0 with $b(\alpha_0) > y_0$, so $\rho(b(\alpha_0)) > b(\alpha_0)$. This is impossible because of the maximizing hypothesis (4.4):

$$\frac{1}{4} \leq 2\alpha[b(\alpha) - \rho(b(\alpha))] \quad \text{for all} \quad \alpha > 0.$$

<u>4.7 Lemma.</u> $\rho(y) = y/2$ <u>for all</u> y.

<u>Proof</u>: By Lemmas 4.5 and 4.6, $\rho(y)/y$ converges to some number L as $y \to \infty$. If $L = \frac{1}{2}$ then 4.5 and 4.6 imply $\rho(y)/y = \frac{1}{2}$ for all y, which is what the theorem claims. So let us prove $L = \frac{1}{2}$.

We claim there is a sequence $\{\alpha_1, \alpha_2, \ldots\}$ such that $\{b(\alpha_n)\} \to \infty$ and ρ is C^1 at $b(\alpha_n)$ and $\rho'(b(\alpha_n)) \to L$. Let us prove this. If it is false then by 4.4 there is some $\delta > 0$ and some $y = y_0$ such that

(i) $\rho'(b(\alpha)) > L + \delta$ for all α such that $b(\alpha) > y_0$ and ρ is C^1 at $b(\alpha)$.

We make y_0 larger, if necessary, so that $y_0 \notin [c_\ell, d_\ell]$ for all ℓ. Next we will show

(ii) $\rho'(y) \geq L + \delta$ if $y > y_0$ and ρ is C^1 at y:

If $y = b(\alpha)$ then (ii) follows from (i) so we need only prove (ii) for $y \in [c_\ell, d_\ell]$ for some ℓ. Since $\rho'(y) = 2\beta_\ell y$ for $y \in [c_\ell, d_\ell]$ it suffices to prove

(iii) $2\beta_\ell c_\ell \geq L + \delta$.

Since $y_0 \notin [c_\ell, d_\ell]$ and ρ is piecewise C^1 we can choose a sequence $\alpha_n \downarrow \beta_\ell$, so $b(\alpha_n) \uparrow c_\ell$, and such that ρ is C^1 at $b(\alpha_n)$ for all n. Then

(iv) $2\alpha_n b(\alpha_n) = \rho'(b(\alpha_n)) > L + \delta$

and we can take the limit on n in (iv) to get

$$2\beta_\ell c_\ell \geq L + \delta \quad .$$

Thus (ii) follows from (i).

By 4.2.2, ρ is continuous and since (1.9b) it is differentiable except on a discrete closed set, and $\rho(0) = 0$, we have

$$\rho(y) = \int_0^y \rho'(t)dt \quad .$$

This and (ii) imply

$$L = \lim_{y\to\infty} \frac{\rho(y)}{y} = \lim_{y\to\infty} \frac{1}{y} \int_0^{y_0} \rho'(t)dt + \lim_{y\to\infty} \frac{1}{y} \int_{y_0}^{y} \rho'(t)dt$$

$$\geq 0 + \lim_{y\to\infty} \frac{1}{y} \int_{y_0}^{y} (L + \delta)dt = \lim_{y\to\infty} \frac{1}{y} (L + \delta)(y - y_0) = L + \delta \quad ,$$

a contradiction. Since (ii) yields a contradiction, a sequence $\{\alpha_1,\alpha_2,\ldots\}$ such that $\{b(\alpha_n)\} \to \infty$, with $2\alpha_n b(\alpha_n) = \rho'(b(\alpha_n)) \to L$, must exist.

We use this sequence $\{\alpha_1,\alpha_2,\ldots\}$ in the maximizing hypothesis (4.1*)

$$\text{(v)} \quad \frac{1}{4} \le 2\alpha_n[b(\alpha_n) - \rho(b(\alpha_n))] = 2\alpha_n b(\alpha_n)[1 - \frac{\rho(b(\alpha_n))}{b(\alpha_n)}] \quad .$$

Since $\rho(y)/y \to L$ as $y \to \infty$, we know $\rho(b(\alpha_n))/b(\alpha_n) \to L$ as $n \to \infty$. Use this and the fact that $2\alpha_n b(\alpha_n) \to L$ as $n \to \infty$ to conclude from (v) that

$$\frac{1}{4} \le \lim_{n\to\infty} 2\alpha_n b(\alpha_n)[1 - \frac{\rho(b(\alpha_n))}{b(\alpha_n)}] = L[1 - L] \quad .$$

By 1.4 we conclude $L = \frac{1}{2}$.

This completes the proof of 4.2.

5. PROOF OF MAIN THEOREM

5.1 Theorem. Suppose Φ is the set of quadratic disutilities,

$$\Phi = \{\varphi_\alpha(y): \alpha > 0\}$$

where $\varphi_\alpha(y) = \alpha y^2$ for $y \geq 0$, and P is the permissible reward functions (defined in 1.9). Then ρ^* is the unique efficient (as defined in 3.1) reward function with respect to Φ and P.

Proof. To prove the theorem we must show that if ρ is a permissible reward function, $b(\cdot)$ is compatible with ρ, and

$$\text{(i)} \qquad \inf\{e(\rho^*,b^*(\cdot),\alpha): \alpha > 0\} \geq \inf\{e(\rho,b(\cdot),\alpha): \alpha > 0\},$$

where $b^*(\cdot)$ is the unique (by 1.3) optimal selection compatible with ρ^*, then $\rho = \rho^*$. First we claim that, with P = the permissible rewards,

$$\text{(ii)} \qquad \hat{\pi}_p(\varphi_\alpha) = \frac{1}{4\alpha} \qquad \text{for all } \alpha > 0.$$

This was proved in 2.2 for P = rewards satisfying 2.1.1-2.1.4, and to prove it in case P = permissible rewards one uses the same proof, except that one needs only prove that the reward function ρ defined in that proof and pictured in Figure 3 is permissible, a trivial task since that ρ is everywhere C^1.

Combining (ii) with the result from Lemma 1.3 that $\pi(\rho^*,b^*(\cdot),\alpha) = 1/8\alpha$, and the definition (3.1) of efficiency, one sees that (i) is equivalent to the maximizing hypothesis previously denoted (4.1*):

$$\text{(5.2)} \qquad \frac{1}{4} \leq 2\alpha[b(\alpha) - \rho(b(\alpha))] \qquad \text{for all } \alpha > 0.$$

Thus, to prove Theorem 5.1 we must show that (5.2) implies $\rho = \rho^*$.

To complete the proof of 5.1 we will study ρ and $b(\cdot)$, assumed to satisfy (5.2) for the remainder of this section, and construct a closely related function σ which will also satisfy (5.2) with the same $b(\cdot)$. Then we will invoke Theorem 4.2 to conclude $\sigma = \rho^*$, and this in turn will imply $\rho = \rho^*$.

5.3 Discontinuities of $b(\cdot)$: Definitions of β_ℓ, c_ℓ, d_ℓ

Since the optimal selection $b(\cdot)$ is nonincreasing (4.1i), it has at most countably many discontinuities, say, at $\{\beta_\ell: \ell = 1,2,\ldots\}$, and these must be jump discontinuities. (See Figure 5b, for example.) The set $\{\beta_\ell\}$ may be finite or even empty, and we cannot assume $\beta_1 < \beta_2 \ldots$ For each ℓ we define

(5.3.1) $$c_\ell = b(\beta_\ell^+), \quad d_\ell = b(\beta_\ell^-).$$

For each ℓ which does not correspond to a β_ℓ, let $[c_\ell, d_\ell]$ be empty (for example, $c_\ell = 2$, $d_\ell = 1$). Then since b is nonincreasing the intervals $[c_\ell, d_\ell]$ are disjoint. We list for future reference three properties which follow from these definitions, properties of monotone functions, and Lemma 4.1.

(5.3.2) $c_\ell > 0$ for all ℓ (use 4.1ii and 4.1iii to prove this).

(5.3.3) If $c_\ell \le b(\alpha) \le d_\ell$, then $\alpha = \beta_\ell$.

(5.3.4) If $y > 0$ and y is not in the range of $b(\cdot)$ (i.e., $y \neq b(\alpha)$ for all α), then $y \in [c_\ell, d_\ell]$ for some (use 4.1iii and 4.1iv to prove this).

5.4 Lemma. For each ℓ such that $[c_\ell, d_\ell]$ is nonempty,

$$(5.4*) \quad \rho(c_\ell) - \beta_\ell c_\ell^2 = \rho(d_\ell) - \beta_\ell d_\ell^2 \text{ and}$$

$$(5.4**) \quad \rho(c_\ell^-) = \rho(c_\ell) \text{ and}$$

$$(5.4***) \quad \rho(d_\ell^+) = \rho(d_\ell) \ .$$

Proof: We denote β_ℓ, c_ℓ, d_ℓ by β, c, d respectively for this proof. Choose sequences γ_n, δ_n with $\gamma_n \downarrow \beta$, $\delta_m \uparrow \beta$. Since $c = b(\beta^+)$, $d = b(\beta^-)$ it follows that

$$b(\gamma_n) \uparrow c \quad \text{and} \quad b(\delta_m) \downarrow d \ .$$

From the compatibility of b (3.2), applied to $b(\gamma_n)$ and $b(\delta_m)$:

$$\text{(i)} \quad \rho(b(\delta_m)) - \gamma_n[b(\delta_m)]^2 \le \rho(b(\gamma_n)) - \gamma_n[b(\gamma_n)]^2$$

$$\text{(ii)} \quad \rho(b(\delta_m)) - \delta_m[b(\delta_m)]^2 \ge \rho(b(\gamma_n)) - \delta_m[b(\gamma_n)]^2 \ .$$

Taking the limit as $n \to \infty$ in (i), and as $m \to \infty$ in (ii), we get

$$\text{(i')} \quad \rho(b(\delta_m)) - \beta_\ell[b(\delta_m)]^2 \le \rho(c^-) - \beta c^2$$

$$\text{(ii')} \quad \rho(d^+) - \beta d^2 \ge \rho(b(\gamma_n)) - \beta[b(\gamma_n)]^2 \ .$$

Use the compatibility of b again, for $b(\delta_m)$ and $b(\gamma_n)$:

(iii) $\rho(b(\delta_m)) - \delta_m[b(\delta_m)]^2 \geq \rho(c) - \delta_m c^2$

(iv) $\rho(d) - \gamma_n d^2 \leq \rho(b(\gamma_n)) - \gamma_n[b(\gamma_n)]^2$

Then take limits as $m \to \infty$ in (i') and (iii) to get:

(i") $\rho(d^+) - \beta d^2 \leq \rho(c^-) - \beta c^2$

(iii") $\rho(d^+) - \beta d^2 \geq \rho(c) - \beta c^2$

In 1.9d we assumed $\rho(c^-) \leq \rho(c)$, so (i") and (iii') imply

(v) $\rho(d^+) - \beta d^2 \leq \rho(c^-) - \beta c^2 \leq \rho(c) - \beta c^2 \leq \rho(d^+) - \beta d^2$.

Since the expressions $\rho(d^+) - \beta d^2$ on the extreme left and right of (v) are equal, all the inequalities in (v) are actually equalities. In particular, equality of the middle expressions in (v) shows $\rho(c^-) = \rho(c)$, which is (5.4**). Equality of the expressions on the left of (v) is

(vi) $\rho(d^+) - \beta d^2 = \rho(c^-) - \beta c^2$.

Now once we prove 5.4***, that $\rho(d^+) = \rho(d)$, (vi) will imply 5.4*, so it remains only to prove $\rho(d^+) = \rho(d)$. Take limits as $n \to \infty$ in (ii') and (iv):

(ii") $\rho(d^+) - \beta d^2 \geq \rho(c) - \beta c^2$

(iv') $\rho(d) - \beta d^2 \leq \rho(c) - \beta c^2$

We have assumed $\rho(d^+) \leq \rho(d)$ so, as before, all four expressions in (ii") and (iv') are equal, in particular $\rho(d^+) = \rho(d)$.

5.5 Construction, elementary properties of σ

(5.5.1) $$\sigma(y) = \begin{cases} \beta_\ell y^2 + \rho(d_\ell) - \beta_\ell d_\ell^2 & \text{if } y \in [c_\ell, d_\ell] \\ \rho(y) & \text{otherwise} . \end{cases}$$

One easily computes that $\rho(d_\ell) = \sigma(d_\ell)$ for all ℓ . In fact we also have $\rho(c_\ell) = \sigma(c_\ell)$ for all ℓ . To see this, write (5.4*) as

$$\rho(c_\ell) = \beta_\ell c_\ell^2 + \rho(d_\ell) - \beta_\ell d_\ell^2 \quad .$$

This agrees with the definition of $\sigma(c_\ell)$. By definition of σ , $\sigma = \rho$ outside the intervals $[c_\ell, d_\ell]$. Summarizing, we have constructed σ so that

(5.5.2) If $y \notin (c_\ell, d_\ell)$ for all ℓ then $\sigma(y) = \rho(y)$.

5.6 Proposition. $\sigma(b(\alpha)) = \rho(b(\alpha))$ for all $\alpha > 0$.

Remark: This is the "close relation" mentioned above between ρ and σ : it says that they give the same rewards to workers who behave according to the optimal selection $b(\cdot)$.

<u>Proof</u>: If $b(\alpha) \notin (c_\ell, d_\ell)$ for all ℓ we're done by 5.5.2. So suppose $c_\ell < b(\alpha) < d_\ell$. Then $\alpha = \beta_\ell$, so we must prove $\sigma(b(\beta_\ell)) = \rho(b(\beta_\ell))$. Let δ_m be a sequence with $\delta_m \uparrow \beta_\ell$, so $b(\delta_m) \downarrow d_\ell$. By compatibility (3.2) applied to $b(\delta_m)$ and $y = b(\beta_\ell)$,

$$\rho(b(\delta_m)) - \delta_m[b(\delta_m)]^2 \geq \rho(b(\beta_\ell)) - \delta_m[b(\beta_\ell)]^2 .$$

Now take the limit as $m \to \infty$ and use $\rho(d_\ell^+) = \rho(d_\ell)$ from (5.4***) to get:

$$\text{(i)} \quad \rho(d_\ell) - \beta_\ell d_\ell^2 \geq \rho(b(\beta_\ell)) - \beta_\ell[b(\beta_\ell)]^2 .$$

But by compatibility applied to $b(\beta_\ell)$ and $y = d_\ell$, $\leq$ holds in (i), so

$$\text{(ii)} \quad \rho(d_\ell) - \beta_\ell d_\ell^2 = \rho(b(\beta_\ell)) - \beta_\ell[b(\beta_\ell)]^2 .$$

Now notice that by the definition of σ, for $c_\ell \leq y \leq d_\ell$:

$$\text{(iii)} \quad \sigma(y) - \beta_\ell y^2 = \rho(d_\ell) - \beta_\ell d_\ell^2 .$$

Finally, put $y = b(\beta_\ell)$ in (iii) and apply it to (ii) to get

$$\sigma(b(\beta_\ell)) - \beta_\ell[b(\beta_\ell)]^2 = \rho(b(\beta_\ell)) - \beta_\ell[b(\beta_\ell)]^2 .$$

Now cancel $-\beta_\ell[b(\beta_\ell)]^2$ to complete the proof.

<u>5.7 Proposition</u>: <u>The optimal selection $b(\cdot)$ which is compatible with ρ is also compatible with σ</u>.

5.7

Proof. We must prove that for each $\alpha > 0$,

$$\sigma(b(\alpha)) - \alpha[b(\alpha)]^2 \geq \sigma(y) - \alpha y^2 \quad \text{for} \quad y \geq 0 \quad .$$

By 5.6 we need only prove

$$\text{(i)} \quad \rho(b(\alpha)) - \alpha[b(\alpha)]^2 \geq \sigma(y) - \alpha y^2 \quad \text{for} \quad y \geq 0 \quad .$$

If $y \notin (c_\ell, d_\ell)$ for all ℓ , then $\sigma(y) = \rho(y)$ by 5.5.2 and (i) follows since $b(\cdot)$ is compatible with ρ . So we may assume $c_\ell < y < d_\ell$ for some ℓ . First suppose $\alpha \leq \beta_\ell$. Then since $y < d_\ell$ and $\beta_\ell r^2 - \alpha r^2$ is a nondecreasing function of r ,

$$\text{(ii)} \quad \beta_\ell d_\ell{}^2 - \alpha d_\ell{}^2 \geq \beta_\ell y^2 - \alpha y^2 \text{ , or } \beta_\ell d^2 - \beta_\ell y^2 \geq \alpha d_\ell{}^2 - \alpha y^2 \text{ .}$$

It follows from the construction of σ (5.5.1) that

$$\text{(iii)} \quad \beta_\ell d^2 - \beta_\ell y^2 = \sigma(d_\ell) - \sigma(y) \quad .$$

Combine (ii) and (iii) to get

$$\text{(iv)} \quad \sigma(d_\ell) - \alpha d_\ell^2 \geq \sigma(y) - \alpha y^2$$

Thus it suffices to prove (i) in case $y = d_\ell$, which we have already done. The remaining case, $\alpha \geq \beta_\ell$, is handled similarly, with c_ℓ in place of d_ℓ .

<u>5.8 Proposition:</u> <u>σ is nondecreasing.</u>

<u>Proof</u>: Suppose $y < z$ and assume by way of contradiction that $\sigma(y) > \sigma(z)$. These two inequalities imply

$$(*) \quad \sigma(y) - \alpha y^2 > \sigma(z) - \alpha z^2 \quad .$$

Now z is not in the range of $b(\cdot)$, since if $z = b(\alpha)$ for some α then the fact (5.7) that $b(\alpha) = z$ maximizes $\sigma(r) - \alpha r^2$ with respect to r would contradict (*). Since z is not in the range of $b(\cdot)$ it follows from 5.3.4 that $z \in [c_k, d_k]$ for some k . On $[c_k, d_k]$, σ is increasing by construction so $\sigma(y) > \sigma(z)$ and $y < z$ imply

$$\sigma(y) > \sigma(c_k) \quad \text{and} \quad y < c_k \quad .$$

Since $\{\beta_\ell\}$ is countable we can construct a sequence $\{\gamma_n\}$ with $\gamma_n \downarrow \beta_k$ and $\gamma_n \neq \beta_\ell$ for all ℓ and n . Then $b(\gamma_n) \uparrow c_k$ so

$$\rho(b(\gamma_n)) \to \rho(c_k^-) = \rho(c_k)$$

by 5.4**. Since $\gamma_n \notin \{\beta_\ell\}$, $b(\gamma_n) \notin [c_\ell, d_\ell]$ for all ℓ and by 5.5.2,

$$\sigma(b(\gamma_n)) = \rho(b(\gamma_n)) \to \rho(c_k) = \sigma(c_k) \quad .$$

Thus for sufficiently large n ,

$$y < b(\gamma_n) \quad \text{and} \quad \sigma(y) > \sigma(b(\gamma_n)) \quad .$$

But this is impossible as explained in the first three sentences of the proof (take $z = b(\gamma_n)$) .

5.9 Proposition: σ is continuous.

Proof. Since σ is nondecreasing, $\sigma(y^+)$ and $\sigma(y^-)$ exist for all y, so it suffices to prove $\sigma(y^-) = \sigma(y) = \sigma(y^+)$. We will prove only $\sigma(y^-) = \sigma(y)$, the other equality being similar. Thus we can assume $y > 0$, and since σ is nondecreasing it suffices to prove $\sigma(y^-) \geq \sigma(y)$.

If $y \in (c_\ell, d_\ell]$ for some ℓ, then $\sigma(y^-) = \sigma(y)$ because of the continuity of σ on $(c_\ell, d_\ell]$. So we may assume $y \notin (c_\ell, d_\ell]$ for all ℓ.

Because the intervals $(c_\ell, d_\ell]$ are disjoint and $y \notin (c_\ell, d_\ell]$ for all ℓ, there is a sequence $\{y_n\}$ with $y_n > 0$ for all n, $y_n \uparrow y$ and $y_n \notin (c_\ell, d_\ell]$ for all n and ℓ. Since $y_n > 0$, it follows from 5.3.4 that for each n there is β_n with $y_n = b(\beta_n)$. The sequence $\{\beta_n\}$ is decreasing since $\{y_n\}$ is increasing. The sequence $\{\beta_n\}$ is bounded by 4.1iii, so $\beta_n \downarrow \beta$ for some β . To summarize:

$$\beta_n \downarrow \beta \ , \quad b(\beta_n) \uparrow y \quad .$$

Since $b(\beta_n)$ maximizes $\sigma(r) - \beta_n r^2$ with respect to r by 5.7,

$$(*) \qquad \sigma(b(\beta_n)) - \beta_n[b(\beta_n)]^2 \geq \sigma(y) - \beta_n y^2 \quad .$$

Take the limit on n in (*):

$$\sigma(y^-) - \beta y^2 \geq \sigma(y) - \beta y^2 \ ,$$

so $\sigma(y^-) \geq \sigma(y)$.

5.10 Reviewing the hypotheses of Theorem 4.2

Recall our plan to apply Theorem 4.2 to σ. At this point we have proved that σ satisfies all the hypotheses of Theorem 4.2 except permissibility: We have proved (5.8 and 5.9) that σ is continuous and nondecreasing, and that is hypothesis 4.2.2. Hypotheses 4.2.3 and 4.2.4, properties of the intervals $[c_\ell, d_\ell]$, are covered by the definitions and properties listed in 5.3. The maximality hypothesis 4.1 follows since we have assumed it for ρ and σ = ρ at the points b(α) by 5.6. An unnumbered hypothesis, that σ is compatible with b(•), was shown in 5.7.

Our next goal, then, is to prove that σ is permissible, which is Theorem 5.13. We will see in the proof of 5.13 that the only nontrivial part is proving that σ is C^1 wherever ρ is. That is intuitively reasonable: σ is got from ρ by changing ρ on the intervals $[c_\ell, d_\ell]$ to be whatever quadratic function smoothly connects the points $(c_\ell, \rho(c_\ell))$ and $(d_\ell, \rho(d_\ell))$. Thus, if anything, we have "smoothed out" ρ.

The technical difficulties arise when the endpoints $\{c_\ell, d_\ell : \ell = 1,2,\ldots\}$ which give nonempty intervals, have a cluster point. If we could assume this set of endpoints was a set of isolated points things would be much easier, but that is an unjustifiable assumption. Thus we need to prove two messy lemmas, which will be used in the proof of Theorem 5.13.

5.11 Lemma: Suppose ρ is C^1 at $y > 0$, $\rho(y) = \sigma(y)$, and $y \notin (c_\ell, d_\ell]$ for all ℓ. Then $\sigma'_-(y) = \rho'(y)$. If ρ is C^1 at $y \geq 0$ and $y \notin [c_\ell, d_\ell)$ for all ℓ, then $\sigma'_+(y) = \rho'(y)$.

Proof: We will prove only the first statement of the lemma; the proof of the second is similar. Assume ρ is C^1 at $y > 0$, $\sigma(y) = \rho(y)$ and $y \notin (c_\ell, d_\ell]$ for all ℓ.

Define $D(t) = \dfrac{\sigma(t) - \sigma(y)}{t - y}$. Let $\varepsilon > 0$. To prove $\sigma'_-(y) = \rho'(y)$ we must find $\delta > 0$ such that

$$(*) \qquad 0 < y - z < \delta \text{ implies } |D(z) - \rho'(y)| < \varepsilon .$$

Our δ will have to satisfy four conditions; after each condition we indicate why it is satisfied for all sufficiently small δ.

(i) $0 < y - b(\alpha) < \delta$ implies $|D(b(\alpha)) - \rho'(y)| < \frac{\varepsilon}{2}$: Use $\rho(b(\alpha)) = \sigma(b(\alpha))$ for all α and the fact that ρ is differentiable at y.

(ii) If $|b(\alpha) - y| < \delta$ and $|z - y| < \delta$ then $|b(\alpha) - z| < \frac{\varepsilon}{4\alpha}$: Since b is nonincreasing, $b(\alpha) \to 0$ as $\alpha \to \infty$, and $y > 0$, we conclude that if δ is small enough then for some $\beta > 0$, $|b(\alpha) - y| < \delta$ implies $\alpha > \beta$. Also pick $\delta > \frac{\varepsilon}{8\beta}$. Then $|b(\alpha) - y| < \delta$ and $|z - y| < \delta$ imply $|b(\alpha) - z| \le |b(\alpha) - y| + |y - z| < 2\delta < \frac{\varepsilon}{4\beta} < \frac{\varepsilon}{4\alpha}$.

(iii) ρ is C^1 in $(y - \delta, y)$: Since ρ is C^1 except on a closed discrete set.

(iv) $0 < y - z < \delta$ implies $|\rho'(z) - \rho'(y)| < \frac{\varepsilon}{2}$: Follows from (iii).

Having chosen δ small enough to satisfy (i-iv) we choose it (perhaps) smaller, so that

(v) $y - \delta \notin [c_\ell, d_\ell]$ for all ℓ .

This is possible since $y \notin (c_\ell, d_\ell]$ for all ℓ and the intervals $[c_\ell, d_\ell]$ are disjoint.

Now suppose $0 < y - z < \delta$. If $z = b(\alpha)$ for some α we are done by (i), so we can assume $z \in [c_k, d_k]$ for some k . Denote c_k by c, d_k by d, β_k by β, for the rest of this proof. By (v), $y - \delta < c$, and $d < y$ since $y \notin (c,d]$ is a hypothesis. Thus $[c,d]$ is contained in $(y - \delta, y)$. It remains to prove that

$$(*) \qquad \sup\{|D(z) - \rho'(y)| : c \le z \le d\} < \varepsilon \quad .$$

Since σ is continuous, D is continuous. Thus the supremum in (*) is attained at some point in $[c,d]$, say at z_0.

<u>Case I</u>: $z_0 = c$. Since $y - \delta < c$ we can choose a sequence $\{b(\alpha_n)\}$ with $y - \delta < b(\alpha_n) < c$ for all n and $b(\alpha_n) \to c$. By (i), $|D(b(\alpha_n)) - \rho'(y)| < \frac{\varepsilon}{2}$ so taking the limit on n we see $|D(c) - \rho'(y)| \leq \frac{\varepsilon}{2} < \varepsilon$, and (*) holds.

<u>Case II</u>: $z_0 = d$. Since $d < y$ we can find a sequence $\{b(\alpha_m)\}$ with $d < b(\alpha_m) < y$ and $b(\alpha_m) \to d$. Now argue as in Case I.

<u>Case III</u>: $z_o \in (c,d)$. Since σ is equal to the differentiable function $(\beta z^2 + \text{constant})$ for $z \in (c,d)$ and $|D(z_0) - \rho'(y)| > 0$, $|D(z) - \rho'(y)|$ is differentiable at z_0, and since it has a local maximum at z_0 its derivative at z_0 is 0. This implies $D'(z_0) = 0$. A computation using the definition of D and $\sigma(z) = \beta z^2 + \text{constant}$ implies $D(z_0) = 2\beta z_0$. (More simply, one can argue that at a critical point z_0 of the difference quotient D, the line through $(y,\sigma(y))$ and $(z_0,\sigma(z_0))$ must be tangent to σ at z_0 and therefore has slope $=$ $\sigma'(z_0) = 2\beta z_0$ and this slope is clearly also equal to $D(z_0)$). So it remains to prove $|2\beta z_0 - \rho'(y)| < \varepsilon$. By the triangle inequality,

$$\text{(vi)} \quad |2\beta z_0 - \rho'(y)| \leq |2\beta z_0 - 2\beta b(\beta)| + |2\beta b(\beta) - \rho'(y)| .$$

Recall that $b(\beta) \in [c,d]$, so $[c,d] \subseteq (y - \delta, y)$ implies $|b(\beta) - y| < \delta$. We also have $|z_0 - y| < \delta$, so (ii) implies

$$\text{(vii)} \quad |2\beta z_0 - 2\beta b(\beta)| = 2\beta|z_0 - b(\beta)| < 2\beta(\tfrac{\varepsilon}{4\beta}) = \tfrac{\varepsilon}{2} .$$

By (iii) above, ρ is C^1 at $b(\beta)$, so by 4.1v, $\rho'(b(\beta)) = 2\beta b(\beta)$. Hence, using (iv) and the fact that $0 < y - b(\beta) < \delta$ we get

$$\text{(viii)} \quad |2\beta b(\beta) - \rho'(y)| = |\rho'(b(\beta)) - \rho'(y)| < \frac{\varepsilon}{2} .$$

Combining (vi), (vii) and (viii) yields

$$|2\beta z_0 - \rho'(y)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad .$$

This completes the proof of the lemma.

<u>5.12 Lemma</u>. <u>Suppose ρ is C^1 at $y > 0$ and $y \notin (c_\ell, d_\ell)$ for all ℓ. Then σ is differentiable at y and $\sigma'(y) = \rho'(y)$</u>.

<u>Proof</u>: If $y \notin [c_\ell, d_\ell]$ for all ℓ then lemma 5.11 implies $\sigma'_+(y) = \sigma'_-(y) = \rho'(y)$, so $\sigma'(y)$ exists and equals $\rho'(y)$. Thus we can restrict ourselves to the cases $y = c_\ell$ and $y = d_\ell$. We will consider only the case $y = c_\ell$. From lemma 5.11 we know $\sigma'_-(c_\ell)$ exists and equals $\rho'(c_\ell)$, so we need only prove $\sigma'_+(c_\ell) = \rho'(c_\ell)$. Let $\{\alpha_n\}$ be a sequence with $\alpha_n \downarrow \beta_\ell$, then $b(\alpha_n) \uparrow c_\ell$ since $c_\ell = b(\beta_\ell^+)$. Since ρ is piecewise C^1 we can assume ρ is C^1 at $b(\alpha_n)$ for all n, then by the first order condition 4.1v,

$$\rho'(b(\alpha_n)) = 2\alpha_n b(\alpha_n) \quad .$$

Now take the limit on n, and use the fact that ρ is C^1 at $y = c_\ell$, to get

$$\rho'(c_\ell) = 2\beta_\ell c_\ell .$$

But by the definition of σ, $\sigma'_+(c_\ell) = 2\beta_\ell c_\ell$, so we conclude $\rho'(c_\ell) = \sigma'_+(c_\ell)$, as desired.

<u>5.13 Theorem</u>. <u>The function σ is permissible</u>.

<u>Proof</u>: Parts (c) and (d) of the definition (1.9) of permissibility are vacuous for σ since σ is continuous.

Part (a) of permissibility, that $0 \leq \sigma(y) \leq y$, follows from the corresponding properties for ρ and the construction of σ in 5.5: To prove $\sigma(y) \leq y$ note that by 5.5.2 $\sigma = \rho$ outside the intervals (c_ℓ, d_ℓ), and $\rho(y) \leq y$ for all y, so $\sigma(y) \leq y$ outside all the intervals (c_ℓ, d_ℓ), and in (c_ℓ, d_ℓ) the function $\sigma(y)$ is an upwardly convex quadratic by 5.5.1. To prove $0 \leq \sigma(y)$ for all y, recall that by 5.3.2 $c_\ell > 0$ for all ℓ so by 5.5.2 $\sigma(0) = \rho(0) = 0$, and σ is nondecreasing.

All that remains is to prove that σ is continuously differentiable except on a discrete closed set X, and $0 \notin X$. Since ρ has this property, it suffices to prove: If ρ is C^1 at y then σ is C^1 at y.

If ρ is C^1 at y then certainly σ is differentiable at y. This is proved in lemma 5.11 for $y = 0$, in lemma 5.12 for $y > 0$ and $y \notin (c_\ell, d_\ell)$ for all ℓ, and on (c_ℓ, d_ℓ) σ is equal to a differentiable (even quadratic) function. Since ρ is piecewise C^1 it will suffice to prove that if ρ is C^1 at y then $\sigma'(y^+)$ and $\sigma'(y^-)$ are both defined and equal $\sigma'(y)$. The function σ is clearly C^1 on (c_ℓ, d_ℓ) so we can assume $y \notin (c_\ell, d_\ell)$ for all ℓ.

We will prove $\sigma'(y^-)$ exists and equals $\sigma'(y)$, the proof for $\sigma'(y^+)$ being similar.

If $y = d_\ell$ for some ℓ then, since by construction σ is quadratic on $(c_\ell, d_\ell]$, $\sigma'(y^-) = \sigma'(y)$. So we can assume $y \notin (c_\ell, d_\ell]$ for all ℓ. By lemma 5.11 and the differentiability of σ at y, $\sigma'(y) = \rho'(y)$.

Let $\varepsilon > 0$. We must find $\delta > 0$ such that

(i) if $0 < y - z < \delta$ then $|\sigma'(z) - \rho'(y)| < \varepsilon$.

(This part of the proof is similar to lemma 5.11.) Since ρ is C^1 at y we can find $\delta > 0$ so that

(ii) $0 < y - z < \delta$ implies $|\rho'(z) - \rho'(y)| < \varepsilon$.

We can also choose δ small enough that ρ is C^1 in $(y - \delta, y)$. Suppose $z \in (y - \delta, y)$. If $z \notin (c_\ell, d_\ell)$ for all ℓ then by lemma 5.12 $\sigma'(z) = \rho'(z)$ so (i) follows from (ii). The only case remaining is $z \in (c_\ell, d_\ell)$ for some ℓ. Then $\sigma'(w) = 2\beta_\ell w$ for $w \in [c_\ell, d_\ell]$ so

(iii) $\sigma'(c_\ell) < \sigma'(z) < \sigma'(d_\ell)$.

Since both $\sigma'(c_\ell)$ and $\sigma'(d_\ell)$ are within δ of $\rho'(y)$, (iii) implies $\sigma'(z)$ is also. This completes the proof of the theorem that σ is permissible.

5.14 End of Proof of Main Theorem 5.1: It remains only to prove that $\rho(y) = y/2$ for $y \geq 0$. We have just completed showing that the reward function σ which we constructed from ρ satisfies the hypotheses of Theorem 4.2--see the discussion in 5.10 above. Now we apply that Theorem to σ and conclude that $\sigma(y) = y/2$ for all $y \geq 0$. Since $b(\cdot)$ is compatible with σ, by 1.3 $b(\alpha) = 1/4\alpha$ for all $\alpha > 0$, so every positive number y is in the range of $b(\cdot)$. By 5.6, $\sigma = \rho$ on the range of $b(\cdot)$, so we conclude $\rho(y) = y/2$ for $y > 0$. We have assumed (1.9) that ρ is C^1 at 0, thus $\rho(y) = y/2$ for $y \geq 0$. This completes the proof of Theorem 5.1.

6. MORE GENERAL DISUTILITIES

In this and the next section we revert to the less cumbersome notation of Section 1, making the assumption that, for each $\rho \in P$ and $\varphi \in \Phi$, the indirect utility $u(y) = \rho(y) - \varphi(y)$ has a unique global maximizer $b(\rho,\varphi)$ with respect to $y \geq 0$. Thus, for example, we will use $\pi(\rho,\varphi_\alpha)$ in place of the $\pi(\rho,b(\cdot),\alpha)$ of Sections 3-5. This assumption, uniqueness of optimal output, is made only to simplify proofs. It is not essential in the results of Sections 6 and 7. We borrow one notation from Section 3 which will also simplify matters: the efficiency ratio is denoted $e_P(\rho,\varphi)$:

$$(6.1) \qquad e_P(\rho,\varphi) = \frac{\pi(\rho,\varphi)}{\hat{\pi}_P(\varphi)} .$$

We may drop the P and write $e(\rho,\varphi)$ if P is clear from the context.

In this section we will always assume of the set of disutilities Φ that each $\varphi \in \Phi$ satisfies $\varphi(0) = 0$. This can be done without loss of generality since it merely changes the scale of the utility function u.

The next proposition relies mainly on an "inheritance" property of undominatedness. Undominated is defined in Section 1.

6.1 Proposition. Let P = the permissible (see 1.9) rewards and assume Φ contains the set of quadratics $\{\varphi_\alpha: \alpha > 0\}$. Then ρ^* is undominated (see 1.12) with respect to Φ, P, and the payoff function e_P.

Proof. Suppose, on the contrary, that ρ dominates ρ^* for some permissible ρ. Then, since Φ contains $\{\varphi_\alpha: \alpha > 0\}$, we have in particular

$$e(\rho^*,\varphi_\alpha) \leq e(\rho,\varphi_\alpha) \quad \text{for all } \alpha > 0,$$

which implies

$$\inf\{e(\rho^*,\varphi_\alpha): \alpha > 0\} \leq \inf\{e(\rho,\varphi_\alpha): \alpha > 0\}.$$

This contradicts Theorem 5.1 which says ρ^* is the unique efficient reward.

The same trick does not apply to the efficiency property, however. There the situation is much worse.

6.2 Lemma. If the disutility φ satisfies $\varphi' \geq 0$ on $(0,\infty)$ and $\hat{y}$ maximizes $y - \varphi(y)$ with respect to $y \geq 0$, then $\hat{\pi}_P(\varphi) = \hat{y} - \varphi(\hat{y})$, where P = the permissible rewards.

Proof: This is a slight generalization of lemma 2.2, so we will only outline the proof. We could show $\hat{\pi} \leq b - \varphi(b)$ for some real number b, as in the first paragraph of the proof of 2.2. Since $\hat{y}$ is a maximizer of $y - \varphi(y)$, we have $b - \varphi(b) \leq \hat{y} - \varphi(\hat{y})$. We conclude $\hat{\pi} \leq \hat{y} - \varphi(\hat{y})$. To prove $\hat{y} - \varphi(\hat{y}) \leq \hat{\pi}$ we construct a function similar to the one pictured in the proof of 2.2, using the fact that $\varphi' \geq 0$ on $(0,\infty)$.

6.3 Proposition. There is a set Φ of infinitely differentiable disutility functions containing $\{\varphi_\alpha : \alpha > 0\}$, such that ρ^* is not the unique efficient reward function with respect to this Φ and P = the permissible rewards. In fact, for this Φ, if ρ^* is efficient, then so is every permissible reward function ρ.

Proof: We will begin by exhibiting some functions ξ which, when smoothed out so as to be infinitely differentiable and then added to $\{\varphi_\alpha : \alpha > 0\}$, give a class Φ such that

(i) $\inf\{e(\rho^*,\varphi): \varphi \in \Phi\} = 0$.

Then we will show that this implies the proposition.

Consider functions ξ as in Figure 7

The continuous function ξ is zero from 0 to b , has constant slope between one half and one from b to $\hat{y}$, and has constant slope greater than one from $\hat{y}$ on. It is easily seen that b is the maximizer with respect to y of the utility $u(y) = \rho^*(y) - \xi(y)$ and $\hat{y}$ of $y - \xi(y)$. By lemma 6.2,

$$e(\rho^*,\xi) = \pi/\hat{\pi} = \frac{b - \rho^*(b)}{\hat{y} - \xi(\hat{y})} .$$

Now consider changing ξ by moving $\hat{y}$ to the right while keeping the slope between b and $\hat{y}$ unchanged, and keeping b fixed. Clearly $\hat{y} - \xi(\hat{y})$ increases without bound, while $b - \rho^*(b)$ remains constant, so for such ξ's $e(\rho^*,\xi)$ goes to zero. Let $\{\xi_n: n=1,2,\ldots\}$ be a sequence of such ξ's ,with

(ii) $\lim_{n\to\infty} e(\rho^*,\xi_n) = 0$.

We can smooth each of the ξ_n's slightly at their kinks - see the dotted lines in the figure - to make them infinitely differentiable while retaining the property (ii). Now let Φ denote the union of $\{\xi_n: n=1,2,\ldots\}$ and $\{\varphi_\alpha: \alpha > 0\}$. Since $e(\rho^*,\varphi_\alpha) = \frac{1}{2}$ for all $\alpha > 0$ by 1.3 and 2.2, we have proved (i), that the infimum of efficiency ratios is 0.

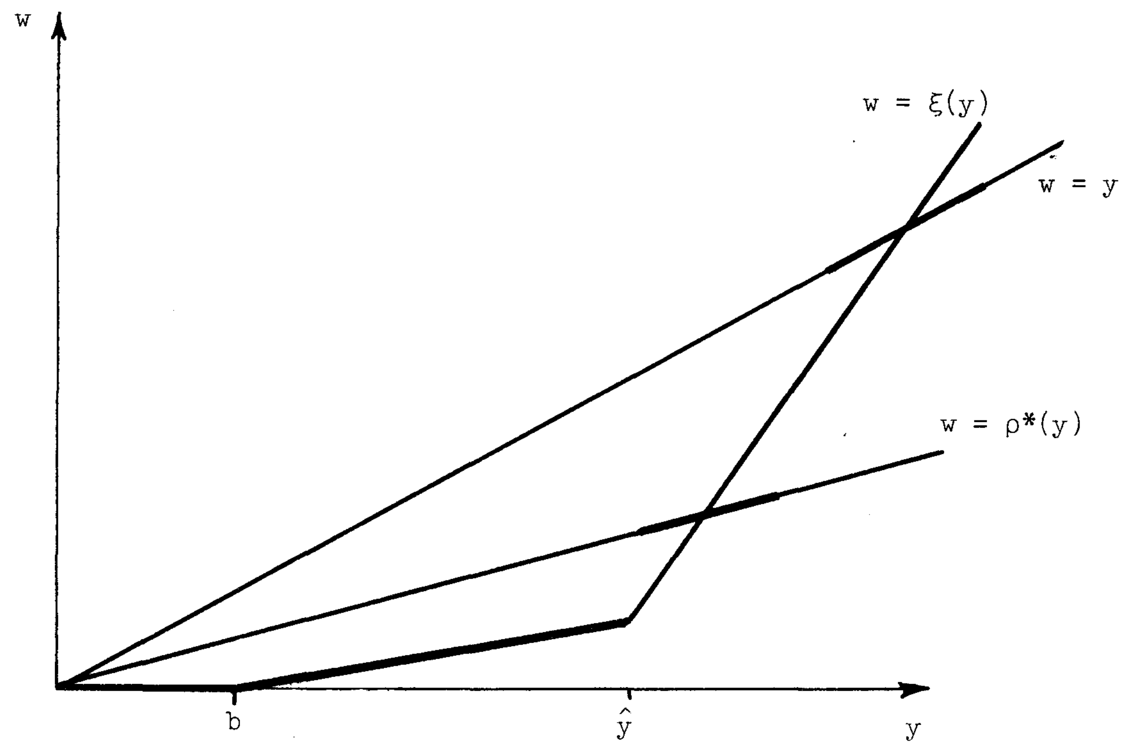


Figure 7

Now suppose ρ^* is efficient for this Φ. Then any other reward function ρ satisfies

(iii) $\inf\{e(\rho,\varphi): \varphi \in \Phi\} \leq 0$

and the reward function ρ is efficient if equality holds in (iii). But for every ρ and every φ, $e(\rho,\varphi) \geq 0$. (To prove this note that we have assumed $\varphi(0) = 0$ for every disutility φ, thus $\hat{y}$ being the maximizer of $y - \varphi(y)$ implies $\hat{\pi} = \hat{y} - \varphi(\hat{y}) \geq 0$ by 6.2, and $\pi = b - \rho(b) \geq 0$ by 1.9a.) Since $e(\rho,\varphi)$ is always non-negative, equality always holds in (iii), so every ρ is efficient, as was claimed. Q.E.D.

Although the unique efficiency property of ρ^* is quite sensitive to the size of Φ (see Section 7), there are some sets Φ for which we can prove that ρ^* remains the unique efficient reward function.

6.4 <u>Proposition.</u> <u>If Φ contains $\{\varphi_\alpha : \alpha > 0\}$, P = permissible rewards, and</u>

$$e(\rho^*,\varphi) \geq \tfrac{1}{2} \text{ for all } \varphi \in \Phi,$$

<u>then ρ^* is the unique permissible efficient reward function with respect to Φ and P.</u>

Proof. If $e(\rho^*,\varphi) \geq \frac{1}{2}$ for all $\varphi \in \Phi$, then since $e(\rho^*,\varphi) = \frac{1}{2}$ for $\alpha > 0$

$$\inf\{e(\rho^*,\varphi): \varphi \in \Phi\} = \tfrac{1}{2} .$$

Now suppose ρ is permissible and efficient, so in particular

$$\inf\{e(\rho,\varphi): \varphi \in \Phi\} \geq \inf\{e(\rho^*,\varphi): \varphi \in \Phi\} = \tfrac{1}{2} .$$

Since $\Phi \supseteq \{\varphi_\alpha : \alpha > 0\}$, this implies

$$\inf\{e(\rho,\varphi_\alpha): \alpha > 0\} \geq \frac{1}{2} = \inf \{e(\rho^*,\varphi_\alpha) : \alpha > 0\}.$$

This says ρ is efficient with respect to $\{\varphi_\alpha : \alpha > 0\}$. Now Theorem 5.1 implies $\rho = \rho^*$. Thus ρ^* is the unique efficient reward function with respect to Φ .

<u>6.5 Which disutilities φ satisfy $e_P(\rho^*,\varphi) \geq \frac{1}{2}$, P = permissible rewards?</u>

The examples ξ in the proof of proposition 6.3 can be altered slightly to show that even if we assumed that $\varphi'(0) = 0$ and that φ' and φ'' are all positive on $(0,\infty)$, it would not follow that $e_P(\rho^*,\varphi) \geq \frac{1}{2}$. But if we assume in addition that the third derivative φ''' is nonnegative then we obtain $e(\rho^*,\varphi) \geq \frac{1}{2}$. We outline a proof of this last assertion using a picture (Figure 8) of the graphs of the derivatives of the functions φ and ρ^*. Note that the assumption $\varphi''' \geq 0$ implies φ' is convex upwards.

The quadratic function $a(y) = \alpha y^2$ is chosen so that its (linear) derivative a' intersects $\rho^{*\prime}$ at the same point as does φ'. (This point is the optimal output b for our φ and for the disutility a, by the appropriate first-order conditions.) Let $\hat{y}_1, \hat{y}_2$ denote the $\hat{y}$ values (see 6.2) for φ and a, respectively. The dotted line in Figure 8 is meant to be tangent to φ' at the point $(b, \frac{1}{2})$. Since $e(\rho^*, a) = \frac{1}{2}$ by 1.3 and 2.2, we need only prove $e(\rho^*, a) \leq e(\rho^*, \varphi)$ in order to show $e(\rho^*, \varphi) \geq \frac{1}{2}$.

By 6.2,

$$e(\rho^*, \varphi) = \frac{b - \rho^*(b)}{\hat{y}_1 - \varphi(\hat{y}_1)} = \frac{\int_0^b (1 - \rho^{*\prime}(t))dt}{\int_0^{\hat{y}_1} (1 - \varphi'(t))dt}$$

and

$$e(\rho^*, a) = \frac{b - \rho^*(b)}{\hat{y}_2 - a(\hat{y}_2)} = \frac{\int_0^b (1 - \rho^{*\prime}(t))dt}{\int_0^{\hat{y}_2} (1 - a'(t))dt} \quad .$$

Since the numerators are equal in these expressions, we need only show

$$\int_0^{\hat{y}_1} (1 - \varphi'(t))dt \leq \int_0^{\hat{y}_2} (1 - a'(t))dt \quad .$$

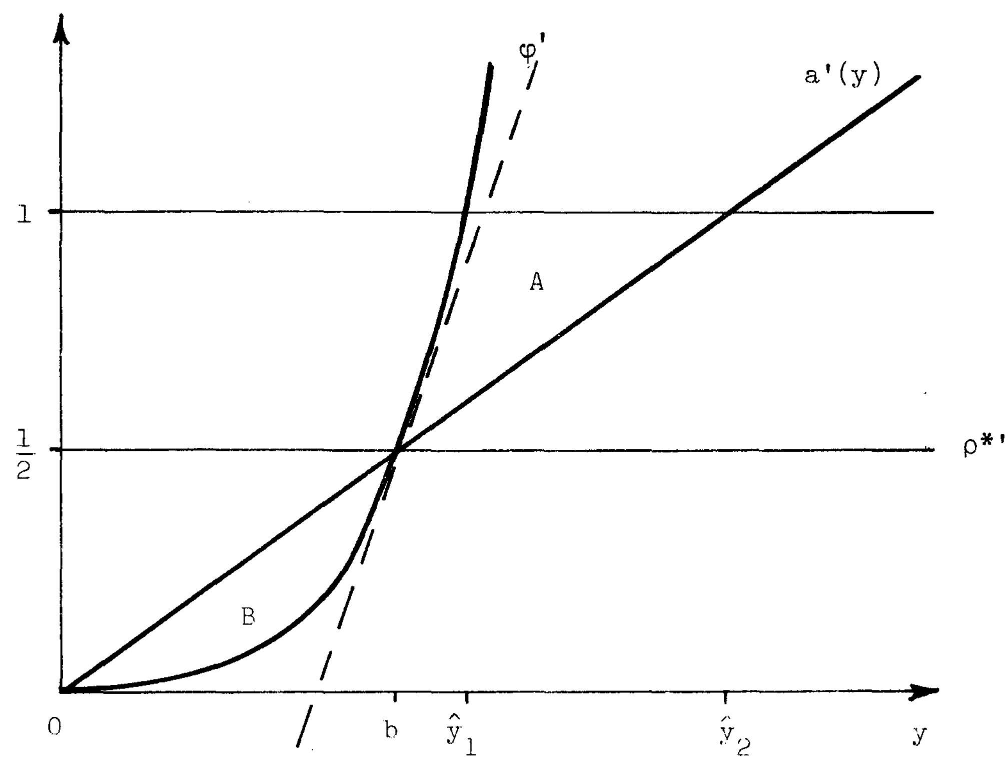


Figure 8

Noticing the areas marked A and B in Figure 8, we see

$$\int_0^{\hat{y}_2}(1-a'(t))dt - \int_0^{\hat{y}_1}(1-\varphi'(t))dt = A - B ,$$

so we need only show $A \geq B$. But this is clear since the two triangles, one formed by a' and the dotted line and $w = 1$, the other by a' and the dotted line and $w = 0$, are of equal area and one contains B , the other is contained in A . We have sketched a proof of:

6.6 Proposition: *If φ is a disutility function satisfying $\varphi'(0) = 0$ and φ', φ'' positive on $(0,\infty)$ and φ''' nonnegative on $(0,\infty)$ then $e_P(\rho^*,\varphi) \geq \frac{1}{2}$, where P = permissible rewards.*

Combining this with 6.4 we obtain:

6.7 Theorem. *If Φ is a set of disutilities containing $\{\varphi_\alpha : \alpha > 0\}$ and each $\varphi \in \Phi$ satisfies $\varphi'(0) = 0$ and φ', φ'' are positive on the interval $(0,\infty)$ and $\varphi''' \geq 0$ on $(0,\infty)$, then ρ^* is the unique permissible efficient reward function with respect to Φ and P = permissible rewards.*

7. FOUR CASES IN WHICH ρ^* IS DOMINATED

In the first two cases, we maintain the assumptions implying that the indirect utility function is given by $u(y) = \rho(y) - \alpha y^2$, but additional information is assumed to be available to the landlord concerning the possible values of α, namely, that they are bounded below (in 7.1) or above (in 7.2) by a known, positive value δ. In both cases, ρ^* is dominated by a permissible reward function (but, by the first sentence of 1.5, ρ^* is not dominated by a linear reward). Of course, without this additional information ρ^* is undominated and even efficient, by Theorem 5.1 and the remark following 1.14.

7.1 Proposition. If $\delta > 0$ and Φ is given by

$$\Phi = \{\varphi_\alpha : \delta < \alpha < \infty\},$$

and P is the permissible reward functions, then the reward function ρ defined by

$$\rho(y) = \frac{y}{2} + \frac{\delta y^2}{2}$$

dominates (see 1.12) ρ^* for the payoff function π and for the payoff function e_P.

Proof. The Proposition will follow if we can prove:

(i) $\pi(\rho^*, \varphi_\alpha) < \pi(\rho, \varphi_\alpha)$ for $\delta < \alpha < \infty$, and

(ii) $\pi(\rho^*, \varphi_\alpha)/\hat{\pi}_P(\varphi_\alpha) < \pi(\rho, \varphi_\alpha)/\hat{\pi}_P(\varphi_\alpha)$ for $\delta < \alpha < \infty$.

Since $\hat{\pi}_P(\varphi_\alpha) > 0$ for all $\alpha > 0$ (see (ii) in the proof of Theorem 5.1), we need only prove (i).

To compute $\pi(\rho, \varphi_\alpha)$ we first compute the maximal output $b(\rho, \varphi_\alpha)$. Since $b(\rho, \varphi_\alpha)$ is the maximizer of the quadratic equation $\rho(y) - \varphi_\alpha(y)$, one easily computes

$$b(\rho,\varphi_\alpha) = \frac{1}{(4\alpha - 2\delta)} .$$

From this, it follows that

(iii) $$\pi(\rho,\varphi_\alpha) = b(\rho,\varphi_\alpha) - \rho(b(\rho,\varphi_\alpha)) = \frac{4\alpha - 3\delta}{2(4\alpha - 2\delta)^2} .$$

Recall from Lemma 1.3 that

(iv) $$\pi(\rho^*,\varphi_\alpha) = \frac{1}{8\alpha} \quad \text{for all } \alpha > 0.$$

Now (i) follows from (iii) and (iv) since

$$\frac{4\alpha - 3\delta}{2(4\alpha - 2\delta)^2} - \frac{1}{8\alpha} = (8\alpha(4\alpha - 2\delta)^2)^{-1}[4\delta(\alpha - \delta)]$$

is positive for $\delta < \alpha$.

The proof of the following proposition is tedious but straightforward, so we omit it.

<u>7.2 Proposition</u>. <u>If Φ is given by</u>

$$\Phi = \{\varphi_\alpha(y): 0 < \alpha < \frac{1}{16\delta}\}$$

<u>and P is the permissible reward functions, then the reward ρ given by</u>

$$\rho(y) = \begin{cases} 0 & \text{for } 0 < y \leq 2\delta \\ \frac{y}{2} - \delta & \text{for } 2\delta < y \end{cases}$$

<u>dominates ρ^* for the payoff function π and for the payoff e_P.</u>

In the third case of this section, the disutility term is assumed to have the special form $\varphi(y) = \alpha y^\beta$, $\alpha > 0$, $\beta > 1$. Now the situation where $\beta \geq 2$[†] and β--like α--is assumed to be unknown to the landlord, is covered

[†]Professor P. N. Bardhan pointed out to us that this is a more appropriate assumption for modeling sharecropping.

by Theorem 6.7, where ρ^* is found to be the unique efficient reward function (hence undominated). On the other hand, when $\beta > 1$ is known to the landlord (while α remains unknown), the linear reward function ρ_k, $k = \frac{1}{\beta}$, is dominant within the class of linear rewards. Hence, when β is known to the landlord, ρ^* is dominated by $\rho_{1/\beta}$ except for $\beta = 2$.

7.3 Proposition. *If $\beta > 1$ and Φ is given by*

$$\Phi = \{\varphi_\alpha^\beta : \alpha > 0\}$$

where $\varphi_\alpha^\beta(y) = \alpha y^\beta$ for all $y \geq 0$, and P is the linear rewards,

$$P = \{\rho_k : k \geq 0\},$$

then the reward function ρ_k with $k = \frac{1}{\beta}$ dominates all other reward functions in P, with respect to P, Φ, and the payoff function π.

Proof. If $\varphi(y) = \alpha y^\beta$, then the maximal output $b(\rho_k, \varphi)$ is a maximizer of

$$\text{(i)} \qquad u(y) = ky - \alpha y^\beta,$$

and, assuming $\beta > 1$, (i) attains its maximum with respect to y at

$$\text{(ii)} \qquad b(\rho_k, \varphi) = \left(\frac{k}{\beta\alpha}\right)^{1/(\beta-1)} .$$

If we plug the value (ii) into the equation for gain, we get

$$\text{(iii)} \qquad \pi(\rho_k, \varphi) = (\beta\alpha)^{-1/(\beta-1)}(1 - k)k^{1/(\beta-1)} ,$$

and the maximum of (iii) with respect to k occurs only at $k = \frac{1}{\beta}$.

The fourth and final case broadens the scope of inquiry in that it involves utility functions that are not linear with respect to reward. (Such functions may be more appropriate to a general equilibrium analysis.)

However, we confine ourselves to a very narrow class of such functions, viz. those of the form $U(r,z) = r^{\gamma} - \psi(z)$, $\gamma \geq 1$; in particular, U remains additively separable with respect to reward versus effort. When γ is known to the landlord and the disutility $\psi(z)$ is αy^2 with α unknown to the landlord, the reward ρ_k with $k = \frac{\gamma}{2}$ is dominant among linear rewards. We remark without proof that if the value of γ is not known, except that $0 \leq \gamma \leq 1$, **so that in computing efficiency we take the infimum of $\pi/\hat{\pi}$ over all** γ, $0 \leq \gamma \leq 1$, as well as over all quadratic disutilities, then the unique efficient reward (when P = linear rewards) is $\rho(y) = y/4$, $y \geq 0$, not ρ^*.

7.4 Proposition. *Suppose we assume the direct utility function U to be of the form*

$$U(r,z) = r^{\gamma} - \psi(z)$$

for some $\gamma \leq 1$, and otherwise retain all the assumptions and definitions of Section 1. If Φ is the set of all quadratic disutilities and P is the set of all linear rewards, then the reward function ρ_k with $k = \frac{\gamma}{2}$ dominates all other reward functions in P, with respect to P, Φ, and the payoff function π.

Proof. The utility function

$$u(y) = (ky)^{\gamma} - \alpha y^2$$

attains its maximum with respect to y (assuming $\gamma \leq 1$) only at

$$b(\rho_k, \varphi_\alpha) = \left(\frac{\gamma}{2\alpha} k^{\gamma}\right)^{1/(2-\gamma)} .$$

For this optimal output, the gain is

$$\text{(i)} \qquad \pi(\rho_k, \varphi_\alpha) = \left(\frac{\gamma}{2\alpha}\right)^{1/(2-\gamma)} (1 - k) k^{\gamma/(2-\gamma)} .$$

The maximum of (i) with respect to k occurs, by the same computation used in the proof of Proposition 7.3, at $k = \frac{\gamma}{2}$. This completes the proof.

BIBLIOGRAPHY

P. Bardhan and T. N. Srinivasan (1971), Crop-sharing Tenancy in Agriculture: A Theoretical and Empirical Analysis, American Economic Review 61: 48.

C. Bell and P. Zusman (1976), A Bargaining Theoretic Approach to Crop-sharing Contracts, American Economic Review 66: 578-588.

H. Chernoff (1954), Rational Selection of Decision Functions, Econometrica 22: 422-443.

S. N. S. Cheung (1969), The Theory of Share Tenancy, Chicago University Press, Chicago.

L. Hurwicz (1977), On the Interaction Between Information and Incentives in Organizations, in Communication and Control in Society, Ed. Klaus Klippendorf, Gordon & Breach, Scientific Publishers, Inc., New York, N.Y.

M. Keren (1969), Effort and Efficiency, Research Report No. 5, Department of Economics, Hebrew University, Jerusalem.

H. Liebenstein (1966), Allocative Efficiency vs. X-efficiency, American Economic Review 56: 392-415.

J. Marschak (1976), Incentives and Information in Organizations, Western Management Sciences Institute, UCLA, preprint.

N. N. Moiseev (1975), Elements of the Theory of Optimal Systems, Moscow.

J. Mirlees (1973), Optimal Incentive Schemes, preliminary version presented at the First Berlin Symposium on Planning, August 22-24, 1973.

L. Savage (1954), The Foundations of Statistics, Wiley, New York.

J. E. Stiglitz (1974), Incentives and Risk in Sharecropping, Review of Economic Studies 41: 219-55.

I. A. Vatel, F. I. Ereshko (1973), Mathematics of Conflict and Collaboration, published by ZNANYE, Moscow (in Russian).