© Journal of Applied Mathematics & Decision Sciences, 3(1), 85–128 (1999)
Reprints Available directly from the Editor. Printed in New Zealand.

ROBUST INSURANCE MECHANISMS AND THE SHADOW PRICES OF INFORMATION CONSTRAINTS

I.V. EVSTIGNEEV evstign@serv2.cemi.rssi.ru
Central Economics and Mathematics Institute, Academy of Sciences of Russia, Nakhimovsky 47, Moscow, 117418, Russia

W.K. KLEIN HANEVELD w.k.klein.haneveld@eco.rug.nl
Department of Economics, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands

L.J. MIRMAN m8h@server1.mail.virginia.edu
Department of Economics, 114 Rouss Hall, University of Virginia, Charlottesville, Va. 22903-3288, USA

Abstract. We consider a risky economic project that may yield either profits or losses, depending on random events. We study an insurance mechanism under which the plan of project implementation maximizing the expected value of profits becomes optimal almost surely. The mechanism is linear in the decision variables, "actuarially fair" and robust to changes in the utility function. The premium and the compensation in the insurance scheme are expressed through dual variables associated with information constraints in the problem of maximization of expected profits. These dual variables are interpreted as the shadow prices of information. Along with the general model, several specialized models are considered in which the insurance mechanism and the shadow prices are examined in detail.

Keywords: Decisions, Risk, Insurance, Value of Information, Convex Stochastic Optimization, Stochastic Lagrange Multipliers

1. Introduction

Problem statement. We consider the following problem of decision-making under uncertainty. A decision-maker (manager) has to carry out a risky project during a time interval $[t_0, t_1]$. At time t_0 he chooses a plan (decision) $x \in X$ which specifies how the project will be realized. The consequences of the decision x depend on random events. At time t_0 the future course of these events cannot be predicted with certainty; however, the probabilities of the possible outcomes are known. By the end of the time period $[t_0, t_1]$, full information about the stochastic factors which might influence the realization of the project is available. Random outcomes may be both favorable and unfavorable. In case of favorable outcome, the project yields profit, while an unfavorable course of events leads to losses.

Suppose that the manager works for an organization or firm which performs a large number of similar projects. If the organization as a whole is risk-neutral, it

would want each of its managers to maximize expected profits. However "if the reward of the manager depends in some measure on his observed profits and if he is a risk-averter, he will wish to play safe by following a course which leads to more predictable profits, even if the expected value is lower. To avoid this outcome, the organization should provide insurance against unfavorable external contingencies..." (Arrow [3]).

Insurance. In the present paper, we describe an insurance mechanism that motivates the manager to make a decision, $\bar{x} \in X$, yielding the maximum expected profit. We consider an insurance scheme under which the decision $\bar{x}$ appears to be optimal for the manager at *almost all* random situations. Suppose that before making a decision the manager computes possible values of his income, assuming one or another particular realization of future random events. Suppose he takes into account the insurance premium he would pay as well as the compensation provided by the insurance mechanism. Having performed these computations for various admissible decisions $x \in X$ and various states of the world, he will conclude that for almost all random outcomes the decision $\bar{x}$ yields the maximum reward. Thus, the manager will practically never regret that he has chosen the plan $\bar{x}$, rather than some other feasible plan.

We assume that X is a convex set and the manager's payoff function is concave in x. Under these conditions and certain technical assumptions, we prove the existence of a *linear* insurance mechanism possessing the above described property (in this mechanism the premium and the compensation are linear functions of x). Having established the general existence theorem, we then apply it to some specialized models. In those models the insurance scheme we deal with is examined in more detail. In particular, we investigate the random variables describing the payoffs with insurance and without it. We find conditions under which the insurance mechanism "stabilizes" the random payoff in the sense of one or another criterion.

Robustness and linearity. Various aspects of the general economic problem considered in this paper have been analyzed by many authors. One can point, for example, to studies of optimal insurance (Arrow [3], Drèze [8], Borch [7]) and of the principal-agent problem (Grossman and Hart [11], Hart and Holmström [13]). For the most part, the models examined in the literature are described in terms of the individual's utility functions of money. We formulate the problem and give a solution to it without using these functions. The study is addressed to economic situations in which there is no reliable information about individual utilities. If such information is available, then one can employ the conventional methods leading to the construction of an optimal insurance system. It can be easily shown, however, that if we wish to deal with the class of all (state-dependent, increasing) utility functions, then the only way to implement the decision $\bar{x}$ by means of a *linear* insurance scheme is to use the mechanism described above. We call this mechanism *robust*, since it is insensitive to utility changes, and since the desired implementation property is retained over the whole class of utility functions.

We concentrate on the class of linear insurance schemes because such schemes are often used in practice, because they are often cheaper to implement and to compute,

and – what is in fact the main idea of this work – because this relatively narrow class of insurance mechanisms turns out to be sufficient to guarantee robustness. We look at the insurance problem from the standpoint of general equilibrium theory, where, as is well known, the class of linear prices is sufficient for reaching an efficient equilibrium state. There is a clear parallelism between our results and this well-known fact.

As regards to robustness, related ideas about the elimination of utilities were outlined by Sondermann [24] in the context of reinsurance in arbitrage-free markets. We consider a different model and focus on different aspects of the insurance problem.

Mathematical background. The mathematical basis for this work is a theory of stochastic optimization problems with information constraints developed by Rockafellar and Wets [21]. Our model, in which the decision $x \in X$ is made without information about the future random events, may be regarded as a special case of the abstract mathematical model analyzed by the above authors. The existence of a robust insurance mechanism follows (up to some technical details) from the Rockafellar-Wets theorem about the existence of Lagrange multipliers removing the information constraints. The premium and the indemnity in the insurance scheme under consideration can be directly expressed through such Lagrange multipliers.

Prices on information. It is natural to expect that the Lagrange multiplier which corresponds to the information constraint in an economic optimization problem gives an economic evaluation of information, just as the Lagrange multiplier removing the resource constraint evaluates this resource. In order to discuss this idea in rigorous terms, we have to have appropriate mathematical structures and methods for specifying and measuring information. It turns out that the conventional methods (e.g. those based on the notion of Shannon's entropy [25] or Blackwell's comparison of experiments [6]) are not quite appropriate for the above purpose. We outline an alternative approach in which information is defined and measured in terms of its property to increase *flexibility* of admissible decision strategies. Using this approach, we show that Lagrange multipliers associated with the information constraints may be regarded as shadow prices on information. Thus, we establish a direct link between such prices and the robust insurance mechanisms.

The idea of possible application of the Rockafellar-Wets results to insurance theory was put forward by E.B. Dynkin (oral remark in the course of a lecture of R.T. Rockafellar). This remark has served as a starting point for this work.

2. The model

Model specification. We turn to a mathematical description of the problem. Let $f(s,x)$ be a function of a random parameter $s \in S$ (S is a measurable space) and of a vector $x \in X$ (X is a subset of R^n). Let $\mu(ds)$ be a probability measure on S representing the distribution of the random parameter s. Assume that the following conditions hold:

(i) For each $s \in S$, the function $f(s,x)$ is continuous and concave in $x \in X$.

(ii) For each $x \in X$, $f(s,x)$ is measurable in $s \in S$.
(iii) There exists a measurable function $q(s)$ such that

$$|f(s,x)| \leq q(s), \ s \in S, \ x \in X,$$

and $E\,q(s) \equiv \int q(s)\mu(ds) < \infty$.

(iv) The set X is convex and compact. The set int$X$ of interior points of X is non-empty.

Here, X is interpreted as the set of decisions, S as the set of states of the world, and $f(s,x)$ as the payoff function. The decision $x \in X$ is made at time t_0, when the value of the random parameter $s \in S$ is not known. At time t_1 there is full information about s. The distribution $\mu(ds)$ of the parameter s does not depend on the decision x. The decision-maker knows this distribution.

It should be noted that, for the basic results in this paper, the assumption int $X \neq \emptyset$, contained in (iv), does not lead to a loss of generality. This assumption is only imposed for sake of convenience of presentation. One can consider the linear manifold spanned on the convex set X and deal with the interior of X in this space (the relative interior of X, see [20], p. 44). In particular, this remark pertains to Theorems 2.1, 3.1, and 3.2 below.

Optimization problem. We consider the following optimization problem.

$$(\mathcal{P}) \text{ Maximize } E\,f(s,x) \ (= \int f(s,x)\mu(ds)) \text{ over } x \in X. \tag{1}$$

By virtue of (i) - (iii), the function $x \to E\,f(s,x)$ is well-defined, concave and continuous in x. Therefore this function achieves its maximum on the compact set X at some point $\bar{x} \in X$. The point $\bar{x}$ represents a decision maximizing the expected payoff.

Insurance scheme. Let $\tilde{p} \in R^n$ be a vector and $p(s)$ $[p : S \to R^n]$ a measurable vector function with finite expectation $E\,|p(s)|$, where $|\cdot|$ stands for the Euclidean norm in R^n. Every such pair $(\tilde{p}, p(\cdot))$ will be called an *insurance scheme*. For any decision $x \in X$, the scalar product $\tilde{p}x$ is the *premium* and $p(s)x$ the *compensation*. The premium is paid at time t_0. The compensation, which depends on s, is paid off at time t_1. The income with insurance equals

$$f(s,x) - \tilde{p}x + p(s)x\ .$$

In the applications we have in mind, components of the vector x represent certain economic quantities, e.g., the amounts of commodities insured under a risky transportation, or the areas of cultivated land (in a model of agricultural insurance), or the amounts of money invested in different parts of the project, etc. Since $\tilde{p}x$ and $p(s)x$ are linear in x, we deal here with a *linear* insurance scheme. Components of the vectors $\tilde{p} = (\tilde{p}_1, \ldots, \tilde{p}_n)$ and $p(s) = (p_1(s), \ldots, p_n(s))$ may be called insurance prices: $\tilde{p}_i$ are the *premium prices* and $p_i(s)$ are the *compensation prices*.

Actuarially fair premium. In this work, we shall consider only those insurance schemes for which

$$\tilde{p} = E\,p(s). \tag{2}$$

Here $E\,p(s)$ is defined as $(E\,p_1(s), \ldots, E\,p_n(s))$. Condition (2) means that the premium is *actuarially fair.*

Assumption (2) is, of course, an idealization. One can consider more realistic insurance schemes satisfying the condition $\tilde{p} = E\,p(s) + \theta$, where θ is a (relatively small) positive number. Insurance mechanisms of this type will be examined in our next paper. Here, we restrict attention to the case $\theta = 0$. This enables us to simplify the form of presentation and to concentrate on the most essential features of the class of mechanisms under consideration.

The central result. We will analyze the insurance schemes described in Theorem 2.1 below.

THEOREM 2.1 *There exists an insurance scheme $(\tilde{p}, p(\cdot))$ such that $\tilde{p} = E\,p(s)$ and for μ-almost all $s \in S$, the inequality*

$$f(s,x) - \tilde{p}x + p(s)x \leq f(s,\bar{x}) - \tilde{p}\bar{x} + p(s)\bar{x} \tag{3}$$

holds for all $x \in X$.

Relation (3) states that under the insurance mechanism $(\tilde{p}, p(\cdot))$ the decision $\bar{x}$ is optimal in almost all random situations. Recall that $\bar{x}$ stands for the decision maximizing $E\,f(s,x)$ over X. It should be noted, however, that if $\bar{x}$ is any decision and $(\tilde{p}, p(\cdot))$ any insurance scheme satisfying (2) and (3), then $\bar{x}$ is necessarily a maximum point of $E\,f(s,x)$. To prove this take the expectations of both sides of (3) and use (2). By employing (2), we also conclude that the expected reward with insurance is the same as without it: $E\,f(s,x) = E(f(s,x) - \tilde{p}x + p(s)x)$ for all $x \in X$.

Some basic properties of $(\tilde{p}, p(\cdot))$. Let $(\tilde{p}, p(\cdot))$ be an insurance scheme possessing properties (2) and (3). Fix any vector $b \in R^n$ and define $\tilde{q} = \tilde{p} + b$, $q(s) = p(s) + b$. Then the insurance scheme $(\tilde{q}, q(\cdot))$ will satisfy (2) and (3) as well. This means, first of all, that the insurance mechanism under consideration is not unique. Furthermore, the vectors $\tilde{p}$ and $p(s)$ are not necessarily positive. The uniqueness and the positivity of the vectors $\tilde{p}$ and $p(s)$ can be established only if these vectors are appropriately normalized. For example, suppose that s takes two values: $s = 0$ (failure) and $s = 1$ (success). Then a natural normalization is given by the condition $p(1) = 0$, which means that in case of success there is no compensation. In specialized models that will be considered in Section 5 this condition together with some additional assumptions guarantees the nonnegativity and the uniqueness of $\tilde{p}$ and $p(s)$. In Section 4, an example will be presented in which $\tilde{p}$ and $p(s)$ are nonnegative but not necessarily unique, even under the above normalization.

It should be noted that negative values of $\tilde{p}_i$ and $p_i(s)$ have an economic meaning as well. For example, components of the vectors $\tilde{p}$ and $p(s)$ may be negative if insurance is combined with a loan (this is one of the oldest forms of insurance used in the marine practice, see Borch [7]).

Suppose $\bar{x} \in \operatorname{int} X$ and the function $f(s,x)$ is differentiable at the point $\bar{x}$. Then inequality (3) yields

$$\tilde{p} - p(s) = f'(s, \bar{x}) \text{ almost surely (a.s.)} , \tag{4}$$

where $f'(s,x)$ is the gradient of $f(s,x)$ with respect to x. Thus, the function $\tilde{p}-p(s)$ is essentially unique: one can change it only on subsets of S having measure zero.

Consider for the moment the one-dimensional case : $n = 1$ and $X \subseteq R^1$. In this case, inequality (4) means that the difference between the premium price and the compensation price is equal a.s. to the *marginal income* $f'(s,\bar{x})$ for the decision $\bar{x}$. In certain models, we may define favorable (resp. unfavorable) random outcomes as those for which the marginal income is positive (resp. negative). Suppose (4) holds for each $s \in S$. Then, we can conclude that the compensation exceeds the premium if and only if the random outcome is unfavorable.

Utility and optimal insurance. Let us discuss the relationships between our approach and the conventional theory of optimal insurance using utility functions (e.g. [7]). Denote by $\mathcal{U}$ the class of all real-valued functions $U(s,r)$ of $s \in S$ and $r \in (-\infty, +\infty)$ satisfying the following conditions: $U(s,r)$ is measurable in s; $U(s,r)$ is continuous and non-decreasing in r; for any function $\zeta : S \to (-\infty, +\infty)$ with $E|\zeta(s)| < \infty$, the expectation $E|U(s,\zeta(s))|$ is finite. Let $(\tilde{p}, p(\cdot))$ be an insurance scheme possessing property (3). Then we have

$$\max_{x \in X} E\, U(s, f(s,x) - \tilde{p}x + p(s)x) = E\, U(s, f(s,\bar{x}) - \tilde{p}\bar{x} + p(s)\bar{x}) \tag{5}$$

for any function $U \in \mathcal{U}$.

Suppose that a function $U(s,\cdot)$, belonging to the class $\mathcal{U}$, represents the decision-maker's (manager's) utility of money at state s. If the goal of the manager is to maximize the expected utility, he can achieve this goal by taking the decision $\bar{x}$, as relation (5) shows. In this sense, the insurance scheme $(\tilde{p}, p(\cdot))$ makes it possible to implement the decision $\bar{x}$ for any utility function in the class $\mathcal{U}$; therefore we call this scheme *robust*. It is important to note that the choice of $\bar{x}$ is optimal regardless of the concavity or convexity of $U(s,\cdot)$ (i.e., regardless of the manager's attitude toward risk).

Conversely, suppose that (5) holds for all $U \in \mathcal{U}$. Then we can apply (5) to any function U of the form $U(s,r) = u(s) \cdot r$, where $u(s)$ is measurable, positive and bounded. We have

$$E\, u(s)[f(s,x) - \tilde{p}x + p(s)x] \leq E\, u(s)[f(s,\bar{x}) - \tilde{p}\bar{x} + p(s)\bar{x}]$$

for each $x \in X$. This implies

$$\mu\{s : f(s,x) - \tilde{p}x + p(s)x \leq f(s,\bar{x}) - \tilde{p}\bar{x} + p(s)\bar{x}\} = 1,\ x \in X.$$

Since $f(s,x) - \tilde{p}x + p(s)x$ is continuous in x, we conclude that with probability one inequality (3) holds for all $x \in X$. Thus, the truth of (5) for all $U \in \mathcal{U}$ is equivalent to the truth of (3).

If we fix some particular utility function $U(s,r)$, then we can construct an insurance scheme which implements the decision $\bar{x}$ and yields, in general, a greater expected utility than the insurance scheme $(\tilde{p}, p(\cdot))$. Consider a simple example.

Let $U(r)$ be a utility function independent of s. Suppose $U(r)$ is strictly concave, increasing and continuous. Assume that X is an interval in $[0,\infty)$ and $\bar{x}$ is a strictly positive number. Let $(\tilde{q}, q(\cdot))$ be an insurance scheme such that

$$q(s) - \tilde{q} = (E\, f(s,\bar{x}) - f(s,\bar{x}))/\bar{x}.$$

Clearly $E\, q(s) - \tilde{q} = 0$. By using Jensen's inequality, we find

$$E\, U(f(s,x) - \tilde{q}x + q(s)x) \leq U(E\, f(s,x)) \leq U(E\, f(s,\bar{x})) =$$

$$E\, U(f(s,\bar{x}) - \tilde{q}\bar{x} + q(s)\bar{x}) \;\; (x \in X),$$

and so $\bar{x}$ maximizes the expectation of utility. Furthermore, we have

$$E\, U(f(s,\bar{x}) - \tilde{q}\bar{x} + q(s)\bar{x}) = U(E\, f(s,\bar{x})) > E\, U(f(s,\bar{x}) - \tilde{p}\bar{x} + p(s)\bar{x}),$$

provided the variance of $f(s,\bar{x}) - \tilde{p}\bar{x} + p(s)\bar{x}$ is strictly positive. Thus, $(\tilde{q}, q(\cdot))$ is strictly more preferable than $(\tilde{p}, p(\cdot))$. However, the insurance scheme $(\tilde{q}, q(\cdot))$ does not necessarily guarantee the implementation of the decision $\bar{x}$ for other *state-dependent* utility functions $U(s,r)$. This follows easily from the equivalence of (3) and (5) (for all $U \in \mathcal{U}$).

Mathematical results related to Theorem 2.1. Theorem 2.1 can be proved by various ways. For example, one can derive it from the following general fact in convex analysis.

Recall that a linear functional b on R^n is called a support functional of a concave function $\psi(x), x \in X\ (\subseteq R^n)$, at the point $x_0 \in X$ if $\psi(x) - \psi(x_0) \leq b(x - x_0)$, $x \in X$. The set of all such functionals is denoted by $\partial_{x_0}\psi(\cdot)$ $(= \partial\psi(x_0))$.

THEOREM 2.2 *Let $f(s,x)$ be a function defined for s in a measurable space S and for x in a convex compact set $X \subseteq R^n$. Let μ be a probability measure on S. Suppose $f(s,x)$ satisfies conditions (i) - (iii). Then we have*

$$\partial_x \int_S f(s,\cdot)\mu(ds) = \int_S \partial_x f(s,\cdot)\mu(ds).$$

Theorem 2.2 follows from a statement in Ioffe and Tihomirov [15], Section 8.3, Theorem 4. (The statement cited deals with the situation when the domain of $f(s,x)$ depends on s and therefore the result has a more complex form.) A version of the above formula involving conditional expectations is proved in Rockafellar and Wets [22]. The earliest reference to results of this type is, apparently, Ioffe and Tihomirov [14].

To obtain Theorem 2.1 as a consequence of Theorem 2.2, observe the following. Since $\bar{x}$ is a point of maximum of $Ef(s,x)$, we have $0 \in \partial_{\bar{x}} \int f(s,\cdot)\mu(ds)$. By virtue of Theorem 2.2, there is an integrable vector function $l(s)$ such that $l(s) \in \partial_{\bar{x}} f(s,\cdot)$ a.s. and $0 = \int l(s)\mu(ds)$. Consider any $\tilde{p}$ and $p(s)$ satisfying $\tilde{p} - p(s) = l(s)$. Then the relations $\tilde{p} = Ep(s)$ and (3) hold, which proves Theorem 2.1.

In the next section we will deduce Theorem 2.1 from other results (dealing, generally, with not necessarily integral functionals). This will yield an independent proof of the theorem and make the paper self-contained. Furthermore, the course of argumentation will point out important relations between robust insurance mechanisms and the shadow prices of information.

3. Insurance prices and Lagrange multipliers for information constraints

Information constraints. In this section we regard the problem $(\mathcal{P})$ as a stochastic optimization problem with an information constraint. This constraint is represented as a linear equation in the space of decision functions. We define $p(\cdot)$ as a Lagrange multiplier (in a function space) associated with this constraint. We show that $(Ep(\cdot), p(\cdot))$ is an insurance scheme satisfying the conditions described in Theorem 2.1.

Denote by $\mathcal{X}$ the class of all measurable mappings $x : S \to R^n$ such that $x(s) \in X$ a.s. We call these mappings *decision functions* (or *strategies*). Define

$$F(x(\cdot)) = E\, f(s, x(s)),\ x(\cdot) \in \mathcal{X}. \tag{6}$$

The problem $(\mathcal{P})$ can be written as

$$(\mathcal{P})\ \text{Maximize } F(x) \text{ over all vectors } x \in X.$$

Together with this problem we consider the following one:

$$(\mathcal{P}_1)\ \text{Maximize } F(x(\cdot)) \text{ over all functions } x(\cdot) \in \mathcal{X}.$$

In the latter problem, the functional $F(\cdot)$ is maximized over all possible strategies $x(s)$. This means that the decision $x(s)$ is made with full information about s. It is clear that in order to solve $(\mathcal{P}_1)$, we have to maximize $f(s, x)$ over X for every fixed s (taking care of the measurability of the solution obtained). In contrast to $(\mathcal{P}_1)$, the problem $(\mathcal{P})$ deals with the maximization of the functional $F(\cdot)$ over all constant (or a.s. constant) functions $x(\cdot) \in \mathcal{X}$, which can be identified with vectors $x \in X$. In this case, the decision x is taken without information about s.

The problem $(\mathcal{P})$ can be obtained from the problem $(\mathcal{P}_1)$ by adding the constraint:

$$x(s) \text{ is constant (a.s.)}. \tag{7}$$

This constraint can be represented in various equivalent forms. Let us write it in the following form

$$x(s) = E\, x(\cdot) \text{ (a.s.)}. \tag{8}$$

Clearly (7) implies (8) and vice versa. Thus, the problem $(\mathcal{P})$ is equivalent to the following problem:

$$\begin{aligned}(\mathcal{P}_0)\quad & \text{Maximize } F(x(\cdot)) \text{ over all functions}\\ & x(\cdot) \in \mathcal{X} \text{ satisfying the constraint}\\ & x(s) - E\, x(\cdot) = 0 \text{ (a.s.)}\end{aligned} \tag{9}$$

The sets of solutions of $(\mathcal{P})$ and $(\mathcal{P}_0)$ coincide. Consequently, the decision $\bar{x}$, which we defined in the previous section, is a solution to both of these problems.

The equivalent requirements (7)–(9) express the fact that the decision $x(\cdot)$ is made without information about s, and so these requirements may be called *information constraints*. Equation (9) has the form $B[x(\cdot)] = 0$ (a.s.), where B is the linear operator transforming a function $x(s)$ into the function $x(s) - E\,x(\cdot)$. Thus, (9) is a linear operator constraint. Theorem 3.1 below shows that there exists a *Lagrange multiplier* $p(\cdot)$ removing this constraint.

As in Section 2, conditions (i)–(iv) are assumed to hold.

THEOREM 3.1 *There exists a measurable function $p : S \to R^n$ such that*

$$E|p(s)| < \infty \tag{10}$$

and

$$F(x(\cdot)) + E\,p(s)[x(s) - E\,x(\cdot)] \leq F(\bar{x}) \tag{11}$$

for all $x(\cdot) \in \mathcal{X}$.

Let us deduce Theorem 2.1 from Theorem 3.1.

Proof of Theorem 2.1: Consider a function $p(s)$ which satisfies (11) and has finite expectation $E|p(s)|$. Define $\tilde{p} = E\,p(\cdot)$. Since $E\,p(s)[x(s) - E\,x(\cdot)] = E\,x(s)[p(s) - \tilde{p}]$, and $E\,\bar{x}[p(s) - \tilde{p}] = 0$, inequality (11) holds if and only if

$$E\,\varphi(s, x(s)) \leq E\,\varphi(s, \bar{x}),\ x(\cdot) \in \mathcal{X}, \tag{12}$$

where $\varphi(s, x) = f(s, x) - \tilde{p}x + p(s)x$. By using the measurable selection theorem (e.g. see [2], Appendix I), we obtain that there exists a function $\hat{x}(\cdot) \in \mathcal{X}$ possessing the following property: $\varphi(s, \hat{x}(s)) = \max_{x \in X} \varphi(s, x)$ (a.s.). This implies

$$\varphi(s, \hat{x}(s)) \geq \varphi(s, \bar{x}) \text{ (a.s.)}. \tag{13}$$

By combining (12) and (13), we conclude $\varphi(s, \bar{x}) = \varphi(s, \hat{x}(s)) = \max_{x \in X} \varphi(s, x)$ (a.s.), which yields (3). □

REMARK 3.1 In the course of the above proof, we have shown that inequality (11) implies (3). The converse is true as well. To prove this, substitute $x(s)$ into (3) and use the fact that $E\,p(s)\bar{x} - E\,\tilde{p}\bar{x} = 0$, which is a consequence of (2) (according to our agreement, equality (2) is supposed to hold). Thus relations (11) and (3) are equivalent if the functional $F(\cdot)$ is defined by (6). It follows from this remark that if inequality (3) is true for some solution $\bar{x}$ to the problem $(\mathcal{P})$, then (3) is also true for any other solution to $(\mathcal{P})$.

REMARK 3.2 If S consists of a finite number of points, then Theorem 3.1 can easily be obtained by using a standard version of the Kuhn-Tucker theorem for convex optimization problems with linear equality constraints (see, e.g., Ioffe and Tihomirov [15], Section 1.3.2). Under the assumption that S is a Borel subset of R^m, Theorems 3.1 and 2.1 follow from results of Rockafellar and Wets [21], Section

4. Rockafellar and Wets developed a theory of Lagrange multipliers for information constraints, dealing with *multistage* stochastic optimization problems.

A more general result. Theorem 3.1 is a special case of the following, more general, result.

THEOREM 3.2 *Let $F(x(\cdot))$, $x(\cdot) \in \mathcal{X}$, be any concave functional continuous with respect to a.s. convergence. Then for this functional the problem $(\mathcal{P}_0)$ has a solution, $\bar{x}$, and there exists a measurable function $p(s)$, satisfying (10) and (11).*

Recall that $\mathcal{X} = \{x(\cdot) : x(s) \in X \text{ a.s.}\}$, where $X \subseteq R^n$ is a compact convex set with int $X \neq \emptyset$. In the above theorem, it is not assumed that the functional $F(\cdot)$ is representable in the form (6). If $F(x(\cdot)) = E\, f(s, x(s))$, where $f(\cdot,\cdot)$ possesses properties (i)–(iii), then $F(\cdot)$ is concave and continuous with respect to a.s. convergence. Thus, Theorem 3.1 follows from Theorem 3.2. Functionals which are not necessarily representable in the form (6) will be considered in Section 7.

A proof of Theorem 3.2 is contained in a paper by Evstigneev [10], Theorem 1, where an analogous result is established for more general (multistage, discrete time) stochastic optimization problems. See also Back and Pliska [4], where continuous time analogues of the above theorem and their economic applications are discussed. For the reader's convenience, we present a direct proof of Theorem 3.2 in the Appendix.

4. A linear model with two states

Model description. Let us investigate the insurance mechanism under consideration in the simplest possible case: S consists of two points, 0 (failure) and 1 (success); X is a segment $[0, a]$ in the real line; the function $f(s, x)$ is linear in x, i.e.,

$$f(0, x) = -q_0 x\,,\ f(1, x) = q_1 x,$$

where q_0 and q_1 are fixed numbers. We assume that q_0, q_1 and a are strictly positive. In this model, the set X of decisions is a set of real numbers, the segment $[0, a]$. A number $x \in X$ may be interpreted as an "intensity" of realization of the risky project. In accordance with the interpretations mentioned in Section 2, x may represent the amount of money invested in an enterprise, or the amount of commodity shipped under a risky transportation, or the area of cultivated land in a model of agricultural insurance, etc. In case of success, the project yields profit $q_1 x$; failure leads to losses $q_0 x$. The probabilities of the random outcomes $s = 0$ and $s = 1$ are $\pi_0 > 0$ and $\pi_1 > 0$, respectively.

We have

$$F(x) \equiv E\, f(s, x) = -\pi_0 q_0 x + \pi_1 q_1 x, \quad x \in [0, a].$$

This function is linear in x, and so there may be three possible sets of maximum points of $F(x)$ on $[0, a]$: $\{0\}$, $[0, a]$, and $\{a\}$. Of interest for us is the last case, when $F(x)$ has a unique maximum attained at $\bar{x} = a$. This is so if and only if

$$\pi_0 q_0 < \pi_1 q_1. \tag{14}$$

Under condition (14), the maximum expected payoff corresponds to the maximum intensity level of the project: $\bar{x} = a$. Throughout this section, inequality (14) will be supposed to hold.

Characterizing the robust insurance schemes. Let us describe all the insurance schemes $(\tilde{p}, p(\cdot))$ satisfying requirements (2), (3) and the following additional condition

$$p(1) = 0. \tag{15}$$

According to this condition, if the random outcome is favorable, then the compensation equals zero. We have

$$\tilde{p} = E\,p(s) = \pi_0 p(0) + \pi_1 \cdot 0 = \pi_0 p(0), \tag{16}$$

which means that the premium price is equal to the compensation price times the probability of failure. Writing inequality (3) for $s = 0$ and $s = 1$, we find

$$-q_0 x - \tilde{p}x + p(0)x \;\leq\; -q_0 a - \tilde{p}a + p(0)a\,, \tag{17}$$

$$q_1 x - \tilde{p}x \;\leq\; q_1 a - \tilde{p}a\,, \tag{18}$$

$x \in [0, a]$. Recall that $\tilde{p} = \pi_0 p(0)$. In view of this, inequalities (17) and (18) are equivalent to the following ones

$$-q_0 - \pi_0 p(0) + p(0) \;\geq\; 0\,, \tag{19}$$

$$q_1 - \pi_0 p(0) \;\geq\; 0\,. \tag{20}$$

Since $1 - \pi_0 = \pi_1$, we conclude that

$$q_0/\pi_1 \leq p(0) \leq q_1/\pi_0. \tag{21}$$

Consequently, those and only those insurance schemes satisfy conditions (3) and (15) for which we have

$$\tilde{p} = \pi_0 p(0),\; p(1) = 0 \text{ and } p(0) \in [q_0/\pi_1, q_1/\pi_0]. \tag{22}$$

Observe that the interval $[q_0/\pi_1, q_1/\pi_0]$ is non-empty by virtue of assumption (14). Therefore at least one such insurance scheme exists (which also follows of course from the general result, Theorem 2.1).

Inequalities (19) and (20) have a clear economic interpretation. By virtue of (19), $p(0)x \leq q_0 x + \tilde{p}x$, $x \in X$. Thus, in case of failure the compensation $p(0)x$ refunds both the amount of losses, $q_0 x$, and the premium payed, $\tilde{p}x$. In view of (20), we have $\tilde{p}x \leq q_1 x$, $x \in X$, and so the value of the premium does not exceed the value of the profit yielded by the project in the case of success.

Payoff with and without insurance. Consider the random variables

$$\xi(s) = f(s, \bar{x}) - \tilde{p}\bar{x} + p(s)\bar{x} \tag{23}$$

and

$$\eta(s) = f(s, \bar{x}), \tag{24}$$

which specify the random income with and without insurance, respectively (computed for the decision $\bar{x}$). Recall that $\bar{x} = a$, and hence $\eta(s) = -q_0 a$ if $s = 0$ and $\eta(s) = q_1 a$ if $s = 1$. Further, we have

$$\xi(0) = -q_1 a - \pi_0 p(0)a + p(0)a = -q_0 a + \pi_1 p(0)a = (\pi_1 p(0) - q_0)a \tag{25}$$

and

$$\xi(1) = q_1 a - \pi_0 p(0)a = (q_1 - \pi_0 p(0))a. \tag{26}$$

In view of (21), $\xi(0) \geq 0$ and $\xi(1) \geq 0$. Thus, each of the insurance schemes satisfying conditions (22) guarantees that the amount of money left after the realization of the project is nonnegative (the possibility of bankruptcy is excluded).

Three special cases. Let us examine the following three cases: 1) $p(0) = q_0/\pi_1$; 2) $p(0) = q_0 + q_1$; 3) $p(0) = q_1/\pi_0$. In the first case, the compensation $p(0)a$ and the premium $\tilde{p} = \pi_0 p(0)$ take on minimal admissible values (see (21)). In the last case, the premium and the compensation are maximal. The insurance scheme with $p(0) = q_1 + q_0$ has a special property which will be discussed later. Observe that the value $p(0) = q_1 + q_0$ is admissible, since each of the inequalities $q_0/\pi_1 < q_0 + q_1$, $q_0 + q_1 < q_1/\pi_0$ is equivalent to (14).

Case 1 ("normal"). If $p(0) = q_0/\pi_1$, then $\xi(s) = 0$ for $s = 0$ and $(q_1 - q_0\pi_0/\pi_1)a$ for $s = 1$ (see (25) and (26)). It follows from (14) that $\xi(1)$ is strictly positive. Comparing the random variables $\xi(s)$ and $\eta(s)$, we see that $\eta(0) = -q_0 a < 0 = \xi(0) < \xi(1) < \eta(1)$. Thus, insurance increases the smallest value of income and reduces its greatest value. The random variable $\xi(s)$ is "less variable" than $\eta(s)$ in any reasonable sense. In particular, the variance of $\xi = \xi(s)$,

$$\operatorname{Var} \xi = E\xi^2 - (E\xi)^2 = (\pi_0/\pi_1)a^2(\pi_1 q_1 - \pi_0 q_0)^2, \tag{27}$$

is strictly less than the variance of $\eta = \eta(s)$, which is equal to $\pi_0\pi_1 a^2(q_1 + q_0)^2$.

Case 2 ("ideal"). If $p(0) = q_0 + q_1$, then $\xi(0) = \xi(1) = (\pi_1 q_1 - \pi_0 q_0)a = E f(s, \bar{x})$. This means that the value of $\xi(s)$ in the cases of success and failure coincide. Such an insurance system leads to a complete stabilization of income. To achieve this, one has to pay the premium $\tilde{p} = \pi_0(q_0 + q_1)a$ greater than the premium $\tilde{p} = \pi_0(q_0/\pi_1)a$ considered in case 1.

Case 3 ("exotic"). If $p(0) = q_1/\pi_0$, then the premium and the compensation achieve their maximal possible values. In this exotic case, we have $\tilde{p} = q_1$, and so all the profit from the successful realization of the project is spent for the premium. By computing the random variable $\xi(s)$, we obtain $\xi(s) = ((\pi_1/\pi_0)q_1 - q_0)a$ if $s = 0$ and $\xi(s) = 0$ if $s = 1$. It follows that the random income $\xi(s)$ is equal to zero in case of success and is strictly positive in case of failure. Thus, insurance turns the former unfavorable random outcome into a "pleasant surprise". The variance of $\xi(s)$ equals

$$\operatorname{Var} \xi = \pi_0\pi_1 a^2(q_1\pi_1/\pi_0 - q_0)^2 = (\pi_1/\pi_0)a^2(\pi_1 q_1 - \pi_0 q_0)^2, \tag{28}$$

and we can see from (28) that $\operatorname{Var}\xi > \pi_0\pi_1 a^2(q_1+q_0)^2 = \operatorname{Var}\eta$ if $\pi_0 < q_1/2(q_0+q_1)$. Therefore the insurance mechanism with $p(0) = q_1/\pi_0$ increases the variance of income if the probability π_0 is small enough.

This kind of "pathological" phenomena related to insurance is well-known. For example, widely used systems of automobile insurance cannot exclude cases when an indemnity for a stolen old car may exceed its real market value. Further discussion of this topic would lead us to the problems of moral hazard and adverse selection, that are, in general, beyond the scope of the paper (e.g. see Milgrom and Roberts [18]).

5. Nonlinear models with two states

Assumptions. In this section, as in the previous one, we assume that the state space S consists of two points, 1 and 0 (success and failure). The random parameter s takes the values 0 and 1 with probabilities $\pi_0 > 0$ and $\pi_1 > 0$. The set X coincides with the segment $[0, a]$ $(a > 0)$ in the real line. However, in this section the function $f(s, \cdot)$ is not supposed to be linear.

We assume that for each $s \in \{0, 1\}$, the function $f(s, x)$ is continuous and concave in $x \in [0, a]$. Then $f(s, x)$ satisfies conditions (i)-(iii) (see Section 2). Clearly, the set $X = [0, a]$ satisfies (iv).

Let $\bar{x}$ be a point in $[0, a]$ at which the function

$$E\,f(s, x) = \pi_0 f(0, x) + \pi_1 f(1, x) \tag{29}$$

attains its maximum. We impose the following assumptions on $f(\cdot\,,\cdot)$ and $\bar{x}$:

$(f.0)$ For each $s \in S$, we have $f(s, 0) = 0$.

$(f.1)$ The point $\bar{x}$ belongs to the interior $(0, a)$ of the segment $[0, a]$, and the function $f(s, \cdot)$ is differentiable at $x = \bar{x}$ for any $s \in S$.

$(f.2)$ We have $f'(1, \bar{x}) > 0$.

According to $(f.0)$, the decision $x = 0$ (inaction) leads to zero payoff at every state $s \in S$. Assumption $(f.2)$ means that the marginal profit yielded by the plan $\bar{x}$ in case of success is strictly positive. By using $(f.0)$, $(f.2)$ and the concavity of $f(1, \cdot)$, we conclude $f(1, \bar{x}) > 0$. Hence, in case of success the realization of the project at the intensity level $\bar{x}$ gives a strictly positive profit. Writing a necessary condition for an extremum of the function (29) at the point $\bar{x}$, we find

$$\pi_0 f'(0, \bar{x}) + \pi_1 f'(1, \bar{x}) = 0. \tag{30}$$

Consequently,

$$f'(0, \bar{x}) < 0, \tag{31}$$

and so the marginal reward in case of failure is strictly negative. Also, we have $E\,f(s, \bar{x}) \geq 0$, since the function $E\,f(s, x)$ is concave and satisfies $E\,f(s, 0) = 0$ and $\frac{d}{dx}E\,f(s, \bar{x}) = 0$. Thus, the expected profit corresponding to the optimal decision $\bar{x}$ is positive or equal to zero.

We do not assume that the payoff $f(0,\bar{x})$ in case of failure is necessarily negative. Formally, this assumption is not needed for the validity of the results we obtain. However, when we say that in case of failure the implementation of the plan $\bar{x}$ leads to losses, we have in mind the number $f(0,\bar{x})$ is negative. The absolute value of this number, $|f(0,\bar{x})| = -f(0,\bar{x})$, specifies the size of losses.

Formulas for the premium and the compensation. Consider an insurance scheme $(\tilde{p}, p(\cdot))$ for which conditions (2) and (3) are satisfied and the compensation in case of success equals zero: $p(1) = 0$. Rewriting formulas (2) and (3) for the model under consideration, we find:

$$\tilde{p} = \pi_0 p(0); \tag{32}$$

$$f(1,x) - \tilde{p}x \leq f(1,\bar{x}) - \tilde{p}\bar{x},\ x \in [0,a]; \tag{33}$$

$$f(0,x) - \tilde{p}x + p(0)x \leq f(0,\bar{x}) - \tilde{p}\bar{x} + p(0)\bar{x},\ x \in [0,a]. \tag{34}$$

By virtue of $(f.1)$, inequalities (33) and (34) imply

$$\tilde{p} = f'(1,\bar{x}),\ p(0) - \tilde{p} = -f'(0,\bar{x}). \tag{35}$$

In view of (32), we have $p(0) - \tilde{p} = (1-\pi_0)p(0) = \pi_1 p(0)$, and so (35) gives

$$p(0) = -f'(0,\bar{x})/\pi_1. \tag{36}$$

Thus, the insurance mechanism $(\tilde{p}, p(\cdot))$ with properties (32)–(34) is unique and the insurance prices $\tilde{p}$ and $p(\cdot)$ can be computed by using formulas (35) and (36).

According to (35), the premium price $\tilde{p}$ equals the marginal profit $f'(1,\bar{x})$ yielded by the project in case of success. By virtue of (36), the compensation price $p(0)$ is equal to the absolute value $|f(0,\bar{x})| = -f'(0,\bar{x})$ of the marginal losses in case of failure, divided by the probability π_1 of success. In view of $(f.2)$ and (31), we have $\tilde{p} > 0$, $p(0) > 0$, and so

$$\tilde{p}\bar{x} > 0\ ,\ p(0)\bar{x} > 0, \tag{37}$$

since $\bar{x} \in (0,a)$. Setting $x = 0$ in (33), (34) and using $(f.0)$, we find,

$$0 \leq f(1,\bar{x}) - \tilde{p}\bar{x} \tag{38}$$

and

$$0 \leq f(0,\bar{x}) - \tilde{p}\bar{x} + p(0)\bar{x}. \tag{39}$$

The first of these two inequalities shows that the premium $\tilde{p}\bar{x}$ is not greater than the profit $f(1,\bar{x})$ in case of success. By virtue of the second inequality, the compensation $p(0)\bar{x}$ covers both the premium paid and the amount of possible losses, so that the resulting net payoff is nonnegative. Relations (38) and (39) are similar to those we have established in the linear case (see (19) and (20)). Note that the only assumption we used when proving (38) and (39) was $(f.0)$.

In addition to (36), we can derive two more formulas for $p(0)$:

$$p(0) = f'(1, \bar{x}) - f'(0, \bar{x}) \tag{40}$$

and

$$p(0) = f'(1, \bar{x})/\pi_0, \tag{41}$$

which follow easily from (36) and (30).

Comparison with the linear case. To compare the formulas obtained with those in Section 4, we define $q_0 = -f'(0, \bar{x})$, $q_1 = f'(1, \bar{x})$. If the functions $f(0, x)$ and $f(1, x)$ are linear in x, then $f(0, x) = -q_0 x$ and $f(1, x) = q_1 x$. By virtue of (30), we have

$$q_0/\pi_1 = q_1/\pi_0, \tag{42}$$

and so the interval $[q_0/\pi_1, q_1/\pi_0]$ which we considered in Section 4 (see (21)), reduces here to a single point. Formulas (36), (40) and (41) can be rewritten as follows:

$$p(0) = q_0/\pi_1 \ ; \ p(0) = q_0 + q_1 \ ; \ p(0) = q_1/\pi_0. \tag{43}$$

Thus, $p(0)$ satisfies simultaneously all the three equalities corresponding to cases 1–3 in the previous section. This fact is a consequence of relation (42). Recall that in the previous section we had $q_0/\pi_1 < q_1/\pi_0$ instead of (42).

In the linear model with two states, we have established the existence of a whole family of robust insurance schemes differing from each other in their properties. In the nonlinear model under consideration, the robust insurance mechanism defined by (32)-(34) is unique. However, the properties of this mechanism may be quite different for different functions $f(\cdot\,,\cdot)$ and probabilities π_0 and π_1. We shall show this in the course of our further study in the remainder of the present section.

Notions of "variability". We will focus on the comparative analysis of the "variability" (or "riskiness") of the random variables $\xi(s) = f(s, \bar{x}) - \tilde{p}\bar{x} + p(s)\bar{x}$ and $\eta(s) = f(s, \bar{x})$, which describe the random payoffs for the decision $\bar{x}$ with and without insurance, respectively. We first formulate some general facts about ξ and η. We have

$$\xi(0) > \eta(0) \text{ and } \xi(1) < \eta(1). \tag{44}$$

This is so, because $\xi(0) = \eta(0) + (1 - \pi_0)p(0)\bar{x}$ and $\xi(1) = \eta(1) - \tilde{p}\bar{x}$, where $\tilde{p}\bar{x} > 0$ and $p(0)\bar{x} > 0$ in view of (37). Inequalities (44) show that the insurance mechanism under consideration increases the payoff in case of failure and reduces the payoff in case of success. From (38) and (39), we conclude $\xi(0) \geq 0$, $\xi(1) \geq 0$, i.e., the payoff with insurance is always nonnegative. Also, we recall that the expected values of ξ and η coincide: $E\,\xi = E\,\eta$.

We will examine conditions under which one or another of the following relations holds:

$$\xi(0) \ \leq \ \xi(1); \tag{45}$$

$$\xi(0) \ = \ \xi(1); \tag{46}$$

$$|\,\xi(0) - \xi(1)\,| \ \leq \ |\,\eta(0) - \eta(1)\,|. \tag{47}$$

The first of these relations expresses a natural requirement on the insurance mechanism: the payoff $\xi(0)$ in case of failure should not be greater than the payoff $\xi(1)$ in case of success. If (46) holds, then $\xi(s)$ does not depend on s, and so insurance leads to the complete stabilization of the payoff. Inequality (47) is equivalent to the following one: $\mathrm{Var}\xi \leq \mathrm{Var}\,\eta$, since $\mathrm{Var}\xi = \pi_0\pi_1|\,\xi(0) - \xi(1)\,|^2$ and $\mathrm{Var}\,\eta = \pi_0\pi_1|\,\eta(0) - \eta(1)\,|^2$. Thus, (47) is satisfied if and only if the insurance mechanism reduces the variance of the payoff (for the decision $\bar{x}$).

If (45) is true, then by using (44), we find

$$\eta(0) < \xi(0) \leq \xi(1) < \eta(1), \tag{48}$$

which yields (47) with strict inequality:

$$|\,\xi(0) - \xi(1)\,| < |\,\eta(0) - \eta(1)\,|. \tag{49}$$

Observe that properties (48) and $E\,\xi = E\,\eta$ imply

$$E\,u(\xi) \geq E\,u(\eta) \text{ for all concave } u : R^1 \to R^1. \tag{50}$$

Proof of this implication, which is not hard to obtain, is left to the reader. A more general result will be established in the next section (see Theorem 6.1). Property (50) means that every risk-averter would prefer the random payoff ξ to the random payoff η. In this sense, ξ is "less risky" than η (e.g. see Rothschild and Stiglitz [23]).

Conditions for risk reduction. In general, none of relations (45)–(47) follows from hypotheses $(f.0)$–$(f.2)$, which we suppose to hold in the present section. To guarantee the truth of these relations, one has to impose additional conditions on the model. Such conditions are described in Theorem 5.1 below.

Define

$$\psi(x) = f(1,x) - f(0,x)\ ,\ x \in [0,a]. \tag{51}$$

THEOREM 5.1 *(a) Inequality (45) is equivalent to the relation*

$$\psi(\bar{x}) \geq \psi'(\bar{x})\bar{x}. \tag{52}$$

(b) Equality (46) holds if and only if

$$\psi(\bar{x}) = \psi'(\bar{x})\bar{x}. \tag{53}$$

(c) Relation (47) is equivalent to the following one:

$$2\psi(\bar{x}) \geq \psi'(\bar{x})\bar{x}. \tag{54}$$

(d) If the function $\psi(x)$, $x \in [0,\bar{x}]$ is concave, linear or convex, then we have $\xi(0) \leq \xi(1)$, $\xi(0) = \xi(1)$, $\xi(0) \geq \xi(1)$, respectively.

(e) Suppose that $\psi(x) \geq 0$ for $x \in [0, \bar{x}]$. If the function $\sqrt{\psi(x)}$, $x \in [0, \bar{x}]$, is concave, linear, or convex, then the number $|\eta(1) - \eta(0)| - |\xi(1) - \xi(0)|$ is nonnegative, equal to zero, or nonpositive, respectively.

By virtue of (d) and (e), the concavity of $\psi(x)$ is a sufficient condition for (45) and the concavity of $\sqrt{\psi(x)}$ is a sufficient condition for (47). If $\psi(x)$ is linear, then we have (46).

Proof of Theorem 5.1: The proof is based on the formulas

$$\xi(1) - \xi(0) = \psi(\bar{x}) - \psi'(\bar{x})\bar{x}; \tag{55}$$

$$(\eta(1) - \eta(0))^2 - (\xi(1) - \xi(0))^2 = \psi'(\bar{x})\bar{x}[2\psi(\bar{x}) - \psi'(\bar{x})\bar{x}]. \tag{56}$$

To check (55), we write

$$\xi(1) - \xi(0) = f(1, \bar{x}) - f(0, \bar{x}) - p(0)\bar{x} = \psi(\bar{x}) - \psi'(\bar{x})\bar{x},$$

where the last equality holds by virtue of (40). Formula (56) can be deduced from (55) as follows:

$$(\eta(1) - \eta(0))^2 - (\xi(1) - \xi(0))^2 = \psi(\bar{x})^2 - [\psi(\bar{x}) - \psi'(\bar{x})\bar{x}]^2 =$$

$$\psi'(\bar{x})\bar{x}[2\psi(\bar{x}) - \psi'(\bar{x})\bar{x}].$$

Assertions (a) and (b) of the theorem are immediate from (55). Since $\psi'(\bar{x})\bar{x} = p(0)\bar{x} > 0$ (see (37) and (40)), assertion (c) is a consequence of (56). By virtue of $(f.0)$, we have $\psi(0) = 0$. Therefore, the difference $\psi(\bar{x}) - \psi'(\bar{x})\bar{x}$ is nonnegative, nonpositive, or equal to zero if the function $\psi(x)$, $x \in [0, \bar{x}]$, is concave, convex, or linear, respectively. This fact, combined with (56), yields (d). To prove (e), observe that $\psi(\bar{x}) > 0$ because $\psi(x) \geq 0$, $x \in [0, \bar{x}]$, and $\psi'(\bar{x}) > 0$. Consequently, we can write

$$2\psi(\bar{x}) - \psi'(\bar{x})\bar{x} = 2\sqrt{\psi(\bar{x})}(\sqrt{\psi(\bar{x})} - \psi'(\bar{x})\bar{x}/2\sqrt{\psi(\bar{x})}) = 2\Psi(\bar{x})(\Psi(\bar{x}) - \Psi'(\bar{x})\bar{x}),$$

where $\Psi(x) = \sqrt{\psi(x)}$. Thus (e) follows from (56). □

A specific form of the payoff function. Let us apply the results obtained to the model in which the function $f(s, x)$ is given in the form

$$f(s, x) = \begin{cases} h(x) - g(x), & \text{if } s = 1, \\ -g(x), & \text{if } s = 0, \end{cases} \tag{57}$$

where $h(x) \geq 0$ and $g(x) \geq 0$ ($x \in [0, a]$) are the *revenue* and the *cost* functions, respectively. Here, $g(x)$ is the cost of the implementation of the project according to the plan x. In case of failure, the project does not yield any return; the size of losses is $g(x)$, and so we have $f(0, x) = -g(x)$. In case of success, the realization of the plan x enables one to obtain the revenue $h(x)$; thus the net profit $f(1, x)$ equals $h(x) - g(x)$.

We assume that the functions $h(x)$ and $g(x)$ are continuous on $[0, a]$, differentiable on $(0, a)$, and satisfy

$$h(0) = g(0) = 0,\ g'(x) > 0,\ x \in (0, a). \tag{58}$$

The functions $-g(x)$ and $h(x) - g(x)$ are supposed to be concave. We fix a point $\bar{x} \in [0, a]$ which maximizes $E\, f(s, x) = \pi_1 h(x) - g(x)$ and assume that $0 < \bar{x} < a$.

From the above assumptions we conclude $\pi_1 h'(\bar{x}) = g'(\bar{x})$ and $f'(1, \bar{x}) = h'(\bar{x}) - g'(\bar{x}) = \left(\pi_1^{-1} - 1\right) g'(\bar{x}) > 0$. Thus conditions $(f.0) - (f.2)$ are satisfied, and so the results obtained in the previous part of the section can be applied to model (57). In particular, we can write the following formulas expressing $\tilde{p}$ and $p(0)$ through the marginal values of revenues and costs:

$$\tilde{p} = h'(\bar{x}) - g'(\bar{x}), \tag{59}$$

$$p(0) = \frac{g'(\bar{x})}{\pi_1} = h'(\bar{x}) = \frac{h'(\bar{x}) - g'(\bar{x})}{\pi_0}. \tag{60}$$

Equality (59) is a consequence of (35); the equalities in (60) follow from (36), (40) and (41). Observe that $f(0, \bar{x}) = -g(\bar{x}) < 0$ by virtue of (58). We also have $f(1, \bar{x}) = h(\bar{x}) - g(\bar{x}) > 0$.

Risk reduction: conditions on the cost and revenue functions. In the model, where $f(s, x)$ is defined by (57), the function $\psi(x)$ coincides with $h(x)$ (see (51)). Therefore, by using Theorem 5.1, we immediately obtain the following result.

THEOREM 5.2 *If the function $h(x)$ ($x \in [0, \bar{x}]$) is concave, then we have $\xi(0) \leq \xi(1)$. If $h(x)$ is linear, then $\xi(0) = \xi(1)$. If $\sqrt{h(x)}$ is concave, then $|\xi(0) - \xi(1)| \leq |\eta(0) - \eta(1)|$ or, equivalently, Var $\xi \leq$ Var η.*

As we have already noticed, the inequality $\xi(0) \leq \xi(1)$ implies (50). Thus, if the revenue function $h(x)$ is concave, then the insurance mechanism reduces the uncertainty of the payoff in the sense of (50). If the function $h(x)$ is not necessarily concave, but its square root is concave, then insurance reduces the variance of the payoff. However, in this case (50) may fail to hold. If $h(x)$ is linear, then the payoff $\xi(s)$ is non-random and its value is equal to $E\, f(s, \bar{x})$.

Theorem 5.2 gives conditions that guarantee a "regular" behavior of the insurance mechanism in question. Under any of those conditions, we have Var $\xi \leq$ Var η. Let us now consider an example where Var $\xi >$ Var η.

Increasing variance: an example. Fix a real number $m > 3$. Define $\pi_0 = (m+1)^{-1}$, $\pi_1 = m(m+1)^{-1}$, $h(x) = x + x^m$, $g(x) = x^m$, $x \in [0, 2]$. Thus, we have

$$f(s, x) = \begin{cases} x, & \text{if } s = 1, \\ -x^m, & \text{if } s = 0. \end{cases}$$

To find $\bar{x}$, we differentiate the function $E\, f(s, x) = -\pi_0 x^m + \pi_1 x$ and obtain $-\pi_0 m \bar{x}^{m-1} + \pi_1 = 0$. Consequently, $\bar{x} = 1$. The random payoff without insurance is computed by the formulas

$$\eta(0) = f(0, \bar{x}) = -1;\ \eta(1) = f(1, \bar{x}) = 1.$$

By virtue of equalities (59) and (60), we have $\tilde{p} = 1$, $p(0) = h'(\bar{x}) = m + 1$. The payoff with insurance is given by

$$\xi(0) = f(0, \bar{x}) - \tilde{p}\bar{x} + p(0)\bar{x} = m - 1;\ \xi(1) = f(1, \bar{x}) - \tilde{p}\bar{x} = 0.$$

We can see that the income $\xi(1)$ in case of success equals zero, while the income $\xi(0)$ in case of failure is strictly positive. Thus, insurance turns an unfavorable random outcome into a favorable one and vice versa. This situation is similar to that we saw in Section 4 (case 3).

By simple computations, we find $\operatorname{Var}\xi = m(m-1)^2(m+1)^{-2}$ and $\operatorname{Var}\eta = 4m(m+1)^{-2}$. Therefore $\operatorname{Var}\xi > \operatorname{Var}\eta$, since $m > 3$. Moreover, we have $\operatorname{Var}\xi/\operatorname{Var}\eta = (m-1)^2/4 \to \infty$ as $m \to \infty$. Thus, the variance of ξ is greater than the variance of η, and the ratio $\operatorname{Var}\xi/\operatorname{Var}\eta$ may be arbitrarily large.

One can give the following intuitive explanation of the above facts. Suppose that the number m is large. Then so is the derivative $g'(\bar{x})$ of the cost function $g(x) = x^m$ at the point $x = \bar{x}$ (this derivative is precisely equal to m). Consequently, a small reduction of $\bar{x}$ may significantly reduce possible losses. On the other hand, the derivative of the profit function $f(1,x)$ at $x = \bar{x}$ is equal to 1. Thus, a small reduction of $\bar{x}$ cannot essentially reduce the profit in case of success. For this reason, if the decision-maker is risk-averse, then some decision $x < \bar{x}$ may be much more preferable for him than the decision $\bar{x}$. Therefore, a strong additional incentive may be needed to motivate the implementation of the plan $\bar{x}$. The role of such an incentive is played by the large insurance compensation $p(0)\bar{x} = m-1$, which is payed off in case of failure. Once the value of $p(0)\bar{x}$ is large, so are the values of $\xi(0)$ and of $\operatorname{Var}\xi$. On the other hand, the random payoff η does not depend on m: we have $\eta(0) = -1$ and $\eta(1) = 1$. As a consequence of this, the random payoff ξ turns out to be "much more variable" than η for large values of m.

In the above considerations, we have assumed that $m > 3$. Suppose now that $m = 3$. Then we have $\pi_0 = 1/4$, $\pi_1 = 3/4$, $\xi(0) = 2$, $\xi(1) = 0$, $\operatorname{Var}\xi = 3 = \operatorname{Var}\eta$. Here, the variance of ξ is equal to the variance of η, however, (50) fails to hold. Indeed, if $u(r) = \min\{0, 1-r\}$, then $Eu(\xi) = -1/4 < 0 = Eu(\eta)$. This remark shows that the inequality $\operatorname{Var}\xi \leq \operatorname{Var}\eta$ does not imply (50) not only for general random variables ξ and η, but also for those particular random variables ξ and η which arise in our model.

The case $g(x) = x$. The remainder of the section will be devoted to the analysis of a special case of model (57). We shall assume that $g(x) = x$, $x \in [0,a]$, and so

$$f(s,x) = \begin{cases} h(x) - x, & \text{if } s = 1, \\ -x, & \text{if } s = 0. \end{cases} \tag{61}$$

Here, x is interpreted as the amount of money invested in the project. In case of success the project yields the revenue $h(x)$. The net profit $f(1,x)$ equals $h(x) - x$. In case of failure, no revenue is obtained and the amount invested is lost. Thus, we have $f(0,x) = -x$.

Suppose that the function $h(x)$ is concave, continuous on $[0,a]$ and differentiable on $(0,a)$. Furthermore, assume that the following requirements are satisfied:

$$\lim_{x\downarrow 0} h'(x) > \frac{1}{\pi_1}, \quad \lim_{x\uparrow a} h'(x) < \frac{1}{\pi_1}, \tag{62}$$

and $h(0) = 0$. Let $\bar{x}$ be a point in $[0,a]$ maximizing $Ef(s,x) = \pi_1 h(x) - x$. By virtue of (62), we have $\bar{x} \in (0,a)$ and $h'(\bar{x}) = \pi_1^{-1}$.

The above assumptions permit us to use the results which have been established for the payoff function (57). In particular, formulas (59) and (60) can be used. From these formulas, we obtain

$$\tilde{p} = h'(\bar{x}) - 1 = \pi_0/\pi_1,\ p(0) = h'(\bar{x}) = \pi_1^{-1}. \tag{63}$$

Note that in this specialized model, the insurance prices $\tilde{p}$ and $p(0)$ do not depend on the function $h(\cdot)$. The only data needed to compute $\tilde{p}$ and $p(0)$ are the probabilities of failure and success, π_0 and π_1.

By using (63), we find

$$\eta(0) = -\bar{x},\ \eta(1) = h(\bar{x}) - \bar{x},\ \xi(0) = 0,\ \xi(1) = h(\bar{x}) - \bar{x}/\pi_1.$$

Since $\xi(0) = -\bar{x} - \tilde{p}\bar{x} + p(0)\bar{x}$, the equality $\xi(0) = 0$ means that the insurance compensation is exactly equal to the sum of the amount of losses and of the premium paid:

$$p(0)\bar{x} = \bar{x} + \tilde{p}\bar{x}. \tag{64}$$

By virtue of (44) and (38), we have $\eta(1) > \xi(1)$ and $\xi(1) \geq 0$. Moreover, it is easily seen that $\xi(1) > 0$. Indeed, if $\xi(1) = 0$, then $h(\bar{x}) = h'(\bar{x})\bar{x}$. Hence, $h(x) = h'(\bar{x})x = \pi_1^{-1}x$ for any $x \in [0, \bar{x}]$, and so $h'(x) = 1/\pi_1$ for $x \in [0, \bar{x}]$, which contradicts (62). The inequality $\xi(1) > 0$ means that $f(1, \bar{x}) > \tilde{p}\bar{x}$, i.e.,

$$h(\bar{x}) - \bar{x} > \tilde{p}\bar{x} = \pi_0\pi_1^{-1}\bar{x}. \tag{65}$$

From the above considerations, we conclude that the following relations hold: $\eta(0) < \xi(0) = 0 < \xi(1) < \eta(1)$. These relations imply (48), (49) and (50). Note that properties (48)–(50) follow from Theorem 5.2 as well, because the function $h(x)$ is concave by assumption.

Reserves. Model (61) is a convenient vehicle for analyzing the following question. Suppose the project manager is going to make the investment $\bar{x}$. Suppose he has a reserve amount, r, of money which he can use either for a direct compensation of possible losses or for insuring himself against those losses. What are the minimal necessary sizes, r_* and r^*, of the reserve r with insurance and without it?

If the manager is not insured, then clearly his reserve r should not be less than $\bar{x}$, since the amount of losses in case of failure is $\bar{x}$. Consequently, we have that $r^* = \bar{x}$. If the manager wishes to be insured, then he has to pay the premium $\tilde{p}\bar{x} = (\pi_0/\pi_1)\bar{x}$. The premium paid gives a right on the insurance compensation $p(0)\bar{x} = \bar{x} + (\pi_0/\pi_1)\bar{x}$, which covers the losses arising in case of failure. Thus we conclude

$$r_* = \frac{\pi_0}{\pi_1}\bar{x}.$$

The ratio $r_*/r^* = \pi_0/\pi_1$ is small, provided that the probability π_0 of failure is small, which is natural to assume in the present context. Consequently, the insurance mechanism under consideration makes it possible to reduce significantly (π_1/π_0 times) the minimal necessary size of the individual reserve.

Also, it should be noted that there is an important distinction between the reserves r_* and r^*. The former is *renewable*, while the latter is not. The profit $f(1,\bar{x}) = h(\bar{x}) - \bar{x}$, yielded by the project in case of success, exceeds $r_* = \tilde{p}\bar{x}$ (see (65)) and is thus sufficient for renewing the reserve r_*. The insurance compensation, paid off in case of failure, contains not only the amount $\bar{x}$ but also the amount $r_* = \tilde{p}\bar{x}$ (see (64)). Hence, the manager can get back the money he has spent for the premium and use it for the formation of his reserve. Thus the reserve r_* can be renewed both in case of success and in case of failure.

6. Models with general state spaces

The problem of income stabilization. Consider the general model described in Section 2. Assume that the payoff function $f(s,x)$ and the space X ($\subseteq R^n$) of decisions satisfy conditions (i)–(iv). Fix a decision $\bar{x} \in X$ maximizing $E\,f(s,x)$ over $x \in X$. Throughout this section, we shall suppose that $\bar{x} \in \mathrm{int}X$. Furthermore, it will be assumed that for each $s \in S$ the function $f(s,x)$ is differentiable with respect to x at the point $x = \bar{x}$. We shall not impose any restrictions on the measurable space S.

Let $(\tilde{p}, p(\cdot))$ be an insurance scheme satisfying requirements (2) and (3). Consider the random variables $\xi(s) = f(s,\bar{x}) - \tilde{p}\bar{x} + p(s)\bar{x}$ and $\eta(s) = f(s,\bar{x})$ (the random payoffs with insurance and without it). We have

$$\xi(s) = f(s,\bar{x}) - f'(s,\bar{x})\bar{x} \quad \text{(a.s.)}, \tag{66}$$

since $\tilde{p} - p(s) = f'(s,\bar{x})$ a.s. (see (4)). The main theme of this section is a comparative analysis of the random variables $\xi(s)$ and $\eta(s)$. We are interested in conditions under which the random payoff $\xi(s)$ is "less variable" (or "more stable") than $\eta(s)$. We use, basically, two different formalizations of the notion "less variable". To this end, we define appropriate preference relations in the space of random variables $\alpha(s)$, $s \in S$.

By definition, the purpose of the insurance mechanism $(\tilde{p}, p(\cdot))$ is to provide incentives for the realization of the plan (decision) which yields the maximum expected profit. If this plan, $\bar{x}$, is unique, we can say that the insurance mechanism $(\tilde{p}, p(\cdot))$ eliminates the uncertainty in the choice of the plan. Indeed, with insurance the decision $\bar{x}$, and only it, becomes optimal for almost all $s \in S$. Of course this does not mean that the uncertainty in the payoff yielded by the plan $\bar{x}$ can always be eliminated. However, there are natural conditions under which the insurance mechanism $(\tilde{p}, p(\cdot))$ *reduces* this uncertainty, i.e., makes the random payoff more stable in the sense of one or another criterion. In the previous section, we described such conditions for nonlinear models with two states. In the present section, we obtain similar results for more general models. These results enable us to outline a unified approach to the analysis of such aspects of insurance as support of effective decisions under risk and stabilization of the random income.

Two preference relations on the space of random variables. Denote by m the product measure $\mu \times \mu$ on $S \times S$. Consider the space L of random variables

(measurable functions) $\alpha : S \to R^1$. Let $\alpha \in L$ and $\beta \in L$. To formalize the idea that α is less variable than β, we use two basic preference relations, $\preceq_1$ and $\preceq_2$, in the space L.

DEFINITION 6.1 We write $\beta \preceq_1 \alpha$ if

$$|\alpha(s_1) - \alpha(s_2)| \leq |\beta(s_1) - \beta(s_2)| \tag{67}$$

m-almost everywhere on $S \times S$ (m-a.e.).

DEFINITION 6.2 We define $\beta \preceq_2 \alpha$ if $\beta \preceq_1 \alpha$ and, in addition, we have

$$(\alpha(s_1) - \alpha(s_2))(\beta(s_1) - \beta(s_2)) \geq 0 \tag{68}$$

(m-a.e.).

According to Definition 6.1, the random payoff α is preferred to the random payoff β if almost every increment of α is not greater – in its absolute value – than the corresponding increment of β. Definition 6.2 contains an additional requirement that the signs of the increments $\alpha(s_1) - \alpha(s_2)$ and $\beta(s_1) - \beta(s_2)$ can be opposite only for those (s_1, s_2) which belong to a negligible subset of $S \times S$. Note that both relations $\beta \preceq_1 \alpha$ and $\beta \preceq_2 \alpha$ are defined without using expected utilities or any moments of the random variables α and β. (We thus follow our general program: to avoid expected utilities and related concepts in the basic definitions.)

Some conventional risk measures. The theorem below demonstrates links between the preference relations $\preceq_1$and $\preceq_2$ and the traditional measures of risk involving concave utility functions and variances.

THEOREM 6.1 *Let $\alpha \in L$ and $\beta \in L$ be two random variables such that* $-\infty < E\,\alpha, E\,\beta < \infty$.

(a) If $\beta \preceq_1 \alpha$, then $\operatorname{Var}\alpha \leq \operatorname{Var}\beta$.

(b) If $\beta \preceq_2 \alpha$ and $E\alpha = E\beta$, then $E\,u(\alpha) \geq E\,u(\beta)$ for any concave function $u : R^1 \to R^1$.

By virtue of (b), if $\beta \preceq_2 \alpha$, then every risk-averse individual will prefer the random payoff α to the random payoff β. If a weaker relation $\beta \preceq_1 \alpha$ holds, we can only state that $\operatorname{Var}\alpha \leq \operatorname{Var}\beta$.

Observe that the expectations $E\,u(\alpha)$ and $E\,u(\beta)$, appearing in assertion (b), are well-defined. This is so, because for any concave function $u : R^1 \to R^1$ we have $u(r) \leq u(0) + ar$, where a is some real number. The values of $E\,u(\alpha)$ and $E\,u(\beta)$ may be finite or equal to $-\infty$.

Assertion (b) will not be valid if we replace $\preceq_2$ by $\preceq_1$. Indeed, take a random variable α and a concave function u satisfying $E\,\alpha = 0$ and $E\,u(\alpha) < E\,u(-\alpha)$. Define $\beta = -\alpha$. Then $\beta \preceq_1 \alpha$, while $E\,u(\alpha) < E\,u(\beta)$. A similar example was considered in Section 5 (the case $m = 3$).

A necessary condition for $\eta \preceq_1 \xi$. We postpone the proof of Theorem 6.1 till the end of this section. Our main goal is to examine conditions under which one or another of the relations $\eta \preceq_1 \xi$ and $\eta \preceq_2 \xi$ holds. Recall that $\eta(s) = f(s, \bar{x})$, $s \in S$, and $\xi(s) = f(s, \bar{x}) - f'(s, \bar{x})\bar{x}$ a.s. (see (66)). It will be convenient to assume that the last formula for $\xi(s)$ holds for all $s \in S$, rather than a.s.. This will not lead to a loss of generality.

We first formulate a simple condition that is necessary for $\eta \preceq_1 \xi$ (and consequently, for $\eta \preceq_2 \xi$). Define $\psi(s_1, s_2, x) = f(s_1, x) - f(s_2, x)$ $(s_1, s_2 \in S,\ x \in X)$. Denote by $\psi'(s_1, s_2, \bar{x})$ the gradient of the function $\psi(s_1, s_2, \cdot)$ at the point $\bar{x}$. This gradient exists, since $f(s, x)$ is differentiable at $x = \bar{x}$ by assumption.

PROPOSITION 6.1 *If $\eta \preceq_1 \xi$, then we have*

$$\psi(s_1, s_2, \bar{x})\psi'(s_1, s_2, \bar{x})\bar{x} \geq 0 \tag{69}$$

almost everywhere with respect to the measure m.

Proof: By using the definitions of ξ and η, we write

$$\xi(s_1) - \xi(s_2) = [f(s_1, \bar{x}) - f(s_2, \bar{x})] - [f'(s_1, \bar{x})\bar{x} - f'(s_2, \bar{x})\bar{x}] =$$

$$\psi(s_1, s_2, \bar{x}) - \psi'(s_1, s_2, \bar{x})\bar{x} \tag{70}$$

and

$$\eta(s_1) - \eta(s_2) = \psi(s_1, s_2, \bar{x}). \tag{71}$$

Therefore

$$(\eta(s_1) - \eta(s_2))^2 - (\xi(s_1) - \xi(s_2))^2 =$$

$$2\psi(s_1, s_2, \bar{x})\psi'(s_1, s_2, \bar{x})\bar{x} - [\psi'(s_1, s_2, \bar{x})\bar{x}]^2. \tag{72}$$

Since $\eta \preceq_1 \xi$, the left-hand side of (72) is nonnegative m-a.e., consequently,

$$2\psi(s_1, s_2, \bar{x})\psi'(s_1, s_2, \bar{x})\bar{x} \geq [\psi'(s_1, s_2, \bar{x})\bar{x}]^2 \geq 0 \ (m\text{-a.e}). \tag{73}$$

□

"More favorable" and "less favorable" states. In the course of further analysis, we shall use the following definition.

DEFINITION 6.3 Let $s_1 \in S$ and $s_2 \in S$. We say that the random outcome s_1 is *more favorable* than s_2, if

$$f'(s_1, \bar{x})\bar{x} > f'(s_2, \bar{x})\bar{x}. \tag{74}$$

If (74) holds, we write $s_2 \ll s_1$.

To understand the meaning of (74), assume for the moment that all vectors x in X are nonnegative and $0 \in X$. Let us interpret coordinates x_i, $i = 1, 2, \ldots, n$, of the vector $x = (x_1, \ldots, x_n) \in X$ as intensities of realization of different parts (subprojects) of the project in question. For example, x_i may specify the amount of money or resources invested in the ith subproject. Consider the function

$$f_0(s, \lambda) = f(s, \lambda\bar{x}),$$

where λ is a real number. This function is defined, in particular, for all λ's in a neighborhood of 1 (since $\bar{x} \in \operatorname{int} X$). The derivative

$$\frac{\partial}{\partial\lambda} f_0(s,\lambda)|_{\lambda=1} = f'(s,\bar{x})\bar{x} \tag{75}$$

may be called the *marginal efficiency* of the plan $\bar{x}$ at state $s \in S$. The number defined in (75) characterizes the marginal increase of the profit under a marginal *proportional* increase of all the components $\bar{x}_i$ of the intensity vector $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)$.

According to Definition 6.3, the random outcome s_1 is more favorable than the random outcome s_2 if the marginal efficiency $f'(s,\bar{x})\bar{x}$ at state $s = s_1$ is greater than the marginal efficiency $f'(s,\bar{x})\bar{x}$ at state $s = s_2$. We have

$$p(s_1)\bar{x} - p(s_2)\bar{x} = f'(s_2,\bar{x})\bar{x} - f'(s_1,\bar{x})\bar{x}. \tag{76}$$

Consequently, *the relation $s_2 \ll s_1$ holds if and only if the insurance compensation $p(s_2)\bar{x}$ is greater than the insurance compensation $p(s_1)\bar{x}$.*

Suppose $\eta \preceq_1 \xi$. Consider the set Δ_0 of those pairs (s_1, s_2) which satisfy $f'(s_1,\bar{x})\bar{x} > f'(s_2,\bar{x})\bar{x}$, $f(s_1,\bar{x}) < f(s_2,\bar{x})$. In view of Proposition 6.1, we have that $m(\Delta_0) = 0$. Therefore, if $s_2 \ll s_1$, then the profit $f(s_2,\bar{x})$ at state s_2 can practically never exceed the profit $f(s_1,\bar{x})$ at state s_1. This means that *the criterion of marginal efficiency is compatible with the criterion of profit.* This fact has been obtained under the assumption $\eta \preceq_1 \xi$. Consequently the above assertion is also valid in those models, where $\eta \preceq_2 \xi$.

A necessary and sufficient condition for $\eta \preceq_2 \xi$. Proposition 6.2 which we formulate below contains a condition that is necessary and sufficient for the truth of the relation $\eta \preceq_2 \xi$.

PROPOSITION 6.2 *The relation $\eta \preceq_2 \xi$ holds if and only if we have*

$$m\{(s_1,s_2) : \ \xi(s_1) < \xi(s_2), \ s_2 \ll s_1\} = 0. \tag{77}$$

The idea of condition (77) is as follows: by using the insurance mechanism, one cannot get a greater profit in a less favorable state. If state s_2 is less favorable than state s_1 $(s_2 \ll s_1)$, then the profit $\xi(s_2)$ cannot be greater than the profit $\xi(s_1)$. According to (77), this requirement may be violated only for those pairs (s_1, s_2) which constitute a negligible subset of $S \times S$.

Proof of Proposition 6.2: Observe that (77) is fulfilled if and only if we have

$$(\xi(s_1) - \xi(s_2))(f'(s_1,\bar{x})\bar{x} - f'(s_2,\bar{x})\bar{x}) \geq 0 \tag{78}$$

(m-a.e.). By using (70) and the formula

$$f'(s_1,\bar{x})\bar{x} - f'(s_2,\bar{x})\bar{x} = \psi'(s_1,s_2,\bar{x})\bar{x}, \tag{79}$$

we can rewrite (78) in the form

$$(\psi - \psi')\psi' \geq 0 \tag{80}$$

(m-a.e.), where $\psi = \psi(s_1,s_2,\bar{x})$ and $\psi' = \psi'(s_1,s_2,\bar{x})\bar{x}$. In view of (72), the relation $\eta \preceq_1 \xi$ is equivalent to the inequality

$$2\psi\psi' \geq (\psi')^2 \tag{81}$$

(m-a.e.). By definition, we have $\eta \preceq_2 \xi$ if and only if $\eta \preceq_1 \xi$ and

$$(\xi(s_1) - \xi(s_2))(\eta(s_1) - \eta(s_2)) \geq 0 \tag{82}$$

(m-a.e.). The latter condition can be written as

$$(\psi - \psi')\psi \geq 0. \tag{83}$$

It remains to observe that for any real numbers ψ and ψ', the system of inequalities (81) and (83) is equivalent to inequality (80). □

Conditions for income stabilization. Define

$$\psi_0(s_1, s_2, \lambda) = \psi(s_1, s_2, \lambda\bar{x}) \tag{84}$$

($\lambda \in R^1$, $\lambda\bar{x} \in X$) and

$$\Sigma = \{(s_1, s_2) \in S \times S : \ s_2 \ll s_1\}. \tag{85}$$

THEOREM 6.2 *(a) We have $\eta \preceq_2 \xi$ if and only if*

$$\psi(s_1, s_2, \bar{x}) \geq \psi'(s_1, s_2, \bar{x})\bar{x} \tag{86}$$

m-almost everywhere on the set Σ.

(b) Equality

$$\xi(s) = E\,\xi(s) \text{ (a.s.)} \tag{87}$$

holds if and only if

$$\psi(s_1, s_2, \bar{x}) = \psi'(s_1, s_2, \bar{x})\bar{x} \tag{88}$$

m-almost everywhere on $S \times S$.

(c) Relation $\eta \preceq_1 \xi$ is satisfied if and only if

$$2\psi(s_1, s_2, \bar{x}) \geq \psi'(s_1, s_2, \bar{x})\bar{x} \tag{89}$$

m-almost everywhere on Σ.

Let $0 \in X$ and $\psi(s_1, s_2, 0) = 0$. Then the following assertions are true:

(d) If for m-almost all (s_1, s_2) in Σ the function $\psi_0(s_1, s_2, \lambda)$, $\lambda \in [0, 1]$, is concave with respect to $\lambda \in [0, 1]$, then we have $\eta \preceq_2 \xi$. If this function is linear with respect to $\lambda \in [0, 1]$ for m-almost all (s_1, s_2) in $S \times S$, then equality (87) holds.

(e) If for m-almost all (s_1, s_2) in Σ, we have

$$\psi_0(s_1, s_2, \lambda) \geq 0, \ \lambda \in [0, 1], \tag{90}$$

and the function

$$\sqrt{\psi_0(s_1, s_2, \lambda)}, \ \lambda \in [0, 1],$$

is concave, then $\eta \preceq_1 \xi$.

Discussion of Theorem 6.2. Theorem 6.2 analyzes conditions under which one or another of the relations (**1**) $\eta \preceq_1 \xi$, (**2**) $\eta \preceq_2 \xi$, and (**3**) $\xi = E\xi$ (a.s.) holds. These conditions guarantee a partial (as in cases (**1**) and (**2**)) or complete (as in case (**3**)) stabilization of the payoff. Clearly (**3**) implies (**2**), and (**2**) implies (**1**). By virtue of Theorem 6.1, it follows from (**1**) that $\operatorname{Var}\xi \leq \operatorname{Var}\eta$ and from (**2**) that $E\,u(\eta) \leq E\,u(\xi)$ for any concave $u(\cdot)$. Assertions (a), (b) and (c) yield conditions which are necessary and sufficient for (**2**), (**3**) and (**1**), respectively. Assertions (d) and (e) contain certain sufficient conditions.

By definition, we have $\psi(s_1, s_2, x) = f(s_1, x) - f(s_2, x)$. Consequently, $\psi_0(s_1, s_2, \lambda)$ is equal to $f(s_1, \lambda\bar{x}) - f(s_2, \lambda\bar{x})$. The assumption of concavity of the function $\psi_0(s_1, s_2, \lambda)$, $\lambda \in [0, 1]$, for m-almost all $s_2 \ll s_1$ may be regarded as a version of the hypothesis of decreasing returns to scale. By using this assumption, we prove the relation $\eta \preceq_2 \xi$ (see statement (d) of Theorem 6.2). A stronger condition of linearity of $\psi_0(s_1, s_2, \cdot)$ enables us to obtain the equality $\xi(s) = E\xi(s)$ (a.s.). To establish the relation $\eta \preceq_1 \xi$, it is sufficient to suppose that the function $\sqrt{\psi_0(s_1, s_2, \lambda)}$ is concave for m-almost all $s_2 \ll s_1$. This requirement is weaker than the above hypothesis of decreasing returns to scale. Recall that assertions (d) and (e) are valid under the condition $\psi(s_1, s_2, 0) = 0$ which holds, for example, if $f(s, 0) = 0$, $s \in S$.

Theorem 6.2 generalizes the main content of Theorem 5.1. The latter examines a model with two states, 0 and 1. By virtue of (31), we have $0 \ll 1$. Consequently, $\Sigma = \{(1, 0)\}$ and $\psi(x) = \psi(1, 0, x)$. Also, we note that the inequality $\xi(0) \leq \xi(1)$ is equivalent to $\eta \preceq_2 \xi$ by virtue of Proposition 6.2.

Proof of Theorem 6.2: Define

$$\theta_1(s_1, s_2) = 2\psi(s_1, s_2, \bar{x})\psi'(s_1, s_2, \bar{x})\bar{x} - (\psi'(s_1, s_2, \bar{x})\bar{x})^2, \tag{91}$$

$$\theta_2(s_1, s_2) = [\psi(s_1, s_2, \bar{x}) - \psi'(s_1, s_2, \bar{x})\bar{x}]\psi'(s_1, s_2, \bar{x})\bar{x}. \tag{92}$$

In the course of the proof of Proposition 6.2, we have shown that $\eta \preceq_1 \xi$ if and only if

$$\theta_1(s_1, s_2) \geq 0 \tag{93}$$

m-a.e., and $\eta \preceq_2 \xi$ if and only if

$$\theta_2(s_1, s_2) \geq 0 \tag{94}$$

m-a.e. (see (81) and (80)). For $(s_1, s_2) \in \Sigma$, (93) is equivalent to (89) and (94) is equivalent to (86) because

$$\Sigma = \{(s_1, s_2) : \ \psi'(s_1, s_2, \bar{x})\bar{x} > 0\}. \tag{95}$$

To prove (a) and (c), it is sufficient to show that $\theta_i(s_1, s_2) \geq 0$ m-a.e. if and only if $\theta_i(s_1, s_2) \geq 0$ for m-almost all (s_1, s_2) in Σ $(i = 1, 2)$. This follows from the equality

$$m\{(s_1, s_2) \in S \times S : \ \theta_i(s_1, s_2) < 0\} = \\ 2m\{(s_1, s_2) \in \Sigma : \ \theta_i(s_1, s_2) < 0\}, \tag{96}$$

which, in turn, is a consequence of the relations:

$$\theta_i(s_1, s_2) = \theta_i(s_2, s_1);\ \theta_i(s_1, s_2) = 0 \text{ if } \psi'(s_1, s_2, \bar{x})\bar{x} = 0.$$

Thus, assertions (a) and (c) are proved.

Let (88) hold. Then, by virtue of (70), we have

$$\xi(s_1) - \xi(s_2) = 0 \ (m\text{-a.e.}). \tag{97}$$

Consider the set

$$S_1 = \{s_1 :\ \xi(s_1) = \xi(s_2) \text{ for } \mu\text{-almost all } s_2 \in S\}.$$

It follows from (97) that $\mu(S_1) = 1$. Fix some $s_1 \in S_1$ and define $c = \xi(s_1)$. Then we have $\xi(s) = c$ μ-a.e., which yields (87). Conversely, (87) implies (97), and, in turn, (97) implies (88) by virtue of (70). This proves (b).

For any $(s_1, s_2) \in S \times S$, we can write

$$\begin{aligned}&\psi(s_1, s_2, \bar{x}) - \psi'(s_1, s_2, \bar{x})\bar{x} = \psi_0(s_1, s_2, 1) - \psi_0'(s_1, s_2, 1) = \\ &[\psi_0(s_1, s_2, 1) - \psi_0(s_1, s_2, 0)] - \psi_0'(s_1, s_2, 1),\end{aligned} \tag{98}$$

since $\psi_0(s_1, s_2, 0) = 0$. From (98), we obtain (86) if the function $\psi_0(s_1, s_2, \lambda)$, $\lambda \in [0, 1]$, is concave and (88) if this function is linear. Hence assertion (d) follows from (a) and (b).

To prove (e), fix $(s_1, s_2) \in \Sigma$ such that inequality (90) holds and the function

$$\Psi_0(\lambda) = \sqrt{\psi_0(s_1, s_2, \lambda)},\ \lambda \in [0, 1],$$

is concave. By using (90) and the relations

$$\psi_0'(s_1, s_2, 1) = \psi'(s_1, s_2, \bar{x})\bar{x} > 0,$$

we conclude

$$\psi_0(s_1, s_2, 1) > 0.$$

Therefore

$$\begin{aligned}&2\psi(s_1, s_2, \bar{x}) - \psi'(s_1, s_2, \bar{x})\bar{x} = 2\psi_0(s_1, s_2, 1) - \psi_0'(s_1, s_2, 1) = \\ &2\Psi_0(1)\,(\Psi_0(1) - \Psi_0'(1)) \geq 0,\end{aligned}$$

since $\Psi_0(0) = 0$ and $\Psi_0(\cdot)$ is concave. Thus, (e) is obtained as a consequence of (c). □

Some applications of Theorem 6.2. We apply Theorem 6.2 to a special class of models. Let

$$S = [0, \infty),\ X = [0, a] \subseteq R^1 \tag{99}$$

$(a > 0)$ and

$$f(s,x) = h(s,x) - g(x), \tag{100}$$

where $h(s,x)$ is the *revenue function* (depending on the random parameter s) and $g(x)$ is the *cost function* (independent of s). In accordance with the general assumptions formulated at the beginning of the section, we suppose that the function $f(s,x)$ satisfies (i)–(iii) and is differentiable with respect to x at $x = \bar{x}$ [$\bar{x} \in \operatorname{int} X = (0,a)$]. Furthermore, the function $h(s,x)$ is supposed to be smooth enough: we postulate the existence of all the partial derivatives of $h(s,x)$ which are considered below.

PROPOSITION 6.3 *Let $f(s,0) = 0$. If we have*

$$\frac{\partial^2 h(s,x)}{\partial s\, \partial x} > 0,\ s \in S,\ x = \bar{x}, \tag{101}$$

and

$$\frac{\partial^3 h(s,x)}{\partial s\, \partial x^2} \le 0,\ s \in S,\ x \in [0,\bar{x}], \tag{102}$$

then $\eta \preceq_2 \xi$. If for each $s \in S$ the function $h(s,x)$ is linear with respect to x, then $\xi(s) = \text{const}$ (a.s.).

Proof: It follows from (101) that the derivative $h'(s,\bar{x}) = \frac{\partial}{\partial s} h(s,\bar{x})$ is a strictly increasing function of s. Consequently, the relations $s_2 < s_1$,

$$h'(s_1,\bar{x}) > h'(s_2,\bar{x}),$$

and (74) are equivalent. Therefore $s_2 \ll s_1$ if and only if $s_2 < s_1$, and so

$$\Sigma = \{(s_1,s_2) :\ s_2 < s_1\}$$

(see (85)).

By virtue of (102), the second derivative

$$h''(s,x) = \frac{\partial^2 h(s,x)}{\partial x^2}$$

is a non-increasing function of $s \in [0,\infty)$ for any $x \in [0,\bar{x}]$. Consequently,

$$h''(s_1,x) - h''(s_2,x) \le 0$$

provided $s_1 > s_2$. Hence the function

$$h(s_1,x) - h(s_2,x),\ x \in [0,\bar{x}]$$

is concave for any $s_1 > s_2$. Since

$$\begin{aligned} &\psi_0(s_1,s_2,\lambda) = \psi(s_1,s_2,\lambda\bar{x}) = \\ &f(s_1,\lambda\bar{x}) - f(s_2,\lambda\bar{x}) = h(s_1,\lambda\bar{x}) - h(s_2,\lambda\bar{x}) \end{aligned} \tag{103}$$

($\lambda \in [0,1]$), we conclude that the function $\psi_0(s_1,s_2,\lambda)$, $\lambda \in [0,1]$, is concave for any $(s_1,s_2) \in \Sigma$. By using assertion (d) of Theorem 6.2, we obtain the desired

relation $\eta \preceq_2 \xi$. Thus, the first statement of Proposition 6.3 is proved. The second statement follows from Theorem 6.2 (d) and (103). □

REMARK 6.1 Suppose that $h(s,x)$ is representable in the form

$$h(s,x) = sh_0(x),\ s \in S,\ x \in [0,a], \tag{104}$$

where $h_0(x)$, $x \in [0,a]$, is a twice differentiable function. Then we have

$$\frac{\partial^2 h(s,x)}{\partial s\, \partial x} = h_0'(x),\ \frac{\partial^3 h(s,x)}{\partial s\, \partial x^2} = h_0''(x),$$

and so conditions (101) and (102) are satisfied if $h_0'(\bar{x}) > 0$ and $h_0''(x) \leq 0, \quad x \in [0,\bar{x}]$.

Lemmas needed for proving Theorem 6.1. In the remainder of this section, we prove Theorem 6.1. The proof is based on several auxiliary facts.

LEMMA 6.1 *Let W be a closed subset of $R^d \times R^d$ $(d \geq 1)$ and $\zeta : S \to R^d$ a measurable vector function satisfying*

$$(\zeta(s_1), \zeta(s_2)) \in W \tag{105}$$

for m-almost all (s_1, s_2) in $S \times S$. Then there exists a measurable vector function $\zeta' : S \to R^d$ such that

$$\zeta(s) = \zeta'(s)\ \mu\text{-a.e.} \tag{106}$$

and

$$(\zeta'(s_1), \zeta'(s_2)) \in W \tag{107}$$

for all (s_1, s_2) in $S \times S$.

LEMMA 6.2 *Let $\kappa(dr)$ be a probability measure on R^1 and $e : R^1 \to R^1$ a function satisfying the following requirements*

$$|e(r_1) - e(r_2)| \leq |r_1 - r_2|; \tag{108}$$

$$e(r_1) \leq e(r_2) \quad \text{if } r_1 \leq r_2; \tag{109}$$

$$-\infty < \int_{R^1} e(r)\kappa(dr) = \int_{R^1} r\kappa(dr) < \infty. \tag{110}$$

Then for every concave function $u : R^1 \to R^1$, we have

$$\int_{R^1} u(e(r))\kappa(dr) \geq \int_{R^1} u(r)\kappa(dr). \tag{111}$$

Proofs of the above two lemmas are presented in the Appendix.

LEMMA 6.3 *Let $\alpha \in L$ and $\beta \in L$. If $\beta \preceq_1 \alpha$, then there exist $\alpha', \beta' \in L$ such that*

$$\alpha'(s) = \alpha(s),\ \beta'(s) = \beta(s),\ \mu\text{-a.e.} \tag{112}$$

and

$$|\alpha'(s_1) - \alpha'(s_2)| \leq |\beta'(s_1) - \beta'(s_2)| \tag{113}$$

for all (s_1, s_2) *in* $S \times S$. *If* $\beta \preceq_2 \alpha$ *then there exist* $\alpha', \beta' \in L$ *satisfying (112), (113) and*

$$(\alpha'(s_1) - \alpha'(s_2))(\beta'(s_1) - \beta'(s_2)) \geq 0 \tag{114}$$

for all (s_1, s_2) *in* $S \times S$.

Proof: Define

$$\begin{aligned} W_1 &= \{(a_1, b_1, a_2, b_2) \in R^2 \times R^2 : |a_1 - a_2| \leq |b_1 - b_2|\}; \\ W_2 &= W_1 \cap \{(a_1, b_1, a_2, b_2) : (a_1 - a_2)(b_1 - b_2) \geq 0\}; \\ \zeta(s) &= (\alpha(s), \beta(s)). \end{aligned}$$

We have $\beta \preceq_1 \alpha$ if and only if $(\zeta(s_1), \zeta(s_2)) \in W_1$ and $\beta \preceq_2 \alpha$ if and only if $(\zeta(s_1), \zeta(s_2)) \in W_2$ (m-a.e.). Since W_1 and W_2 are closed, both assertions of Lemma 6.3 follow from Lemma 6.1. □

LEMMA 6.4 *Let* $\alpha, \beta \in L$. *If* $\beta \preceq_1 \alpha$, *then there exists a function* $e : R^1 \to R^1$ *satisfying inequality (108) and such that*

$$\alpha(s) = e(\beta(s)) \ \mu\text{-a.e.}. \tag{115}$$

If $\beta \preceq_2 \alpha$, *then there exists a function* $e : R^1 \to R^1$ *possessing properties (108), (109) and (115).*

Proof: Let us prove the second assertion of the lemma; the first one is proved similarly. By virtue of Lemma 6.3, we may assume without loss of generality that relations (67) and (68) hold for all (s_1, s_2) in $S \times S$, rather than m-a.e.. For each $r \in \beta(S)$, we define $e(r) = \alpha(s)$, where s is any element of $\beta^{-1}(r)$. In view of (67), $\alpha(s)$ does not depend on $s \in \beta^{-1}(r)$, and so $e(r)$ is well-defined for $r \in \beta(S)$. Furthermore, we have

$$e(\beta(s)) = \alpha(s), \ s \in S. \tag{116}$$

It follows from (66) and (67) that

$$0 \leq e(r_2) - e(r_1) \leq r_2 - r_1 \tag{117}$$

for all $r_1, r_2 \in \beta(S)$, $r_1 < r_2$. Indeed, if $r_1 = \beta(s_1) < r_2 = \beta(s_2)$, then $0 \leq \alpha(s_2) - \alpha(s_1) \leq \beta(s_2) - \beta(s_1) = r_2 - r_1$, which yields (117), because $\alpha(s_1) = e(r_1)$ and $e(s_2) = r_2$. Relations (117) mean that the function $e(r)$, $r \in \beta(S)$, satisfies conditions (108) and (109) on the set $\beta(S)$. By using (108), we extend $e(r)$ to the closure $\mathrm{cl}\,\beta(S)$ of the set $\beta(S)$. The extended function $e(r)$ satisfies (108) and (109) on $\mathrm{cl}\,\beta(S)$.

The open set $R^1 \setminus \mathrm{cl}\,\beta(S)$ is a union of a finite or countable family $\mathcal{E}$ of disjoint open intervals whose boundary points belong to $\mathrm{cl}\,\beta(S)$. For a finite interval $\Lambda = (a, b) \in \mathcal{E}$, we define

$$e(r) = \frac{e(b) - e(a)}{b - a}(r - a) + e(a), \quad r \in \Lambda.$$

For an infinite interval $\Lambda = (-\infty, c) \in \mathcal{E}$ or $\Lambda = (c, +\infty) \in \mathcal{E}$, we set

$$e(r) = e(c),\ r \in \Lambda.$$

It is easily seen that the function $e(r)$ $(r \in R^1)$ so defined satisfies (108) and (109). In view of (116), this function satisfies (115). □

Proof of Theorem 6.1: Let $\beta \preceq_1 \alpha$. By using (67), we find

$$\begin{aligned} 2\mathrm{Var}\,\alpha &= \int_S \int_S (\alpha(s_1) - \alpha(s_2))^2 \mu(ds_1)\mu(ds_2) \leq \\ & \int_S \int_S (\beta(s_1) - \beta(s_2))^2 \mu(ds_1)\mu(ds_2) = 2\mathrm{Var}\,\beta, \end{aligned}$$

which proves (a).

Let $\beta \preceq_2 \alpha$. According to Lemma 6.4, there exists a function $e : R^1 \to R^1$ satisfying (115), (108) and (109). Denote by $\kappa(dr)$ the distribution of the random variable $\beta(s)$. We have

$$E\,\beta = \int_{R^1} r\kappa(dr) \quad \text{and} \quad E\,\alpha = \int_{R^1} e(r)\kappa(dr).$$

Since $-\infty < E\,\alpha = E\,\beta < \infty$, we obtain (110). From Lemma 6.2, it follows that

$$E\,u(\alpha) = \int_{R^1} u(e(r))\kappa(dr) \geq \int_{R^1} u(r)\kappa(dr) = E\,u(\beta)$$

for any concave $u(\cdot)$. □

7. The shadow prices of information constraints

The model and the assumptions. In Section 3 we defined a Lagrange multiplier $p(\cdot)$ associated with the information constraint (9) in the optimization problem $(\mathcal{P}_0)$. Using this Lagrange multiplier, we proved the existence of a robust insurance mechanism. Components of the vector function $p(\cdot)$ played the roles of compensation prices in the robust insurance scheme. In the present section, we consider other applications of the Lagrange multipliers removing the information constraints. We use them for a study of the value of information in problems of decision making under risk. The main goal of the study is to develop a concept of shadow prices on information analogous to the well-known concept of shadow prices on resources (e.g. see Birchenhall and Grout [5]).

This section deals with the following model. We are given a space X of decisions and a state space S. It is assumed that the set X is contained in R^n and satisfies conditions formulated in (iv) (see Section 2). The set S is finite: $S = \{s_1, \ldots, s_N\}$. There is a random parameter s taking values in S. This parameter takes each of the values $s_1, \ldots, s_N$ with strictly positive probabilities $\pi_1, \ldots, \pi_N$. We assume that a concave continuous functional $F(x(\cdot))$ is defined on the set $\mathcal{X} = \{x(\cdot) : x(s) \in X, s \in S\}$ of decision functions (strategies) $x(\cdot) : S \to R^n$. The continuity of $F(\cdot)$ means that $F(x_k(\cdot)) \to F(x(\cdot))$ if $x_k(s) \to x(s)$ for each $s \in S$. (Here, convergence of $x_k(s)$ for each s is equivalent to a.s. convergence.)

We consider the optimization problems $(\mathcal{P})$ and $(\mathcal{P}_1)$ described in Section 3. Recall that the problem $(\mathcal{P}_1)$ deals with the maximization of $F(x(\cdot))$ over all functions $x(\cdot) \in \mathcal{X}$. In the problem $(\mathcal{P})$, we maximize $F(x)$ over all constant functions in $\mathcal{X}$, which can be identified with elements x in X. We fix a solution to $(\mathcal{P})$ and denote this solution by $\bar{x}$. Clearly $\bar{x}$ exists since $F(x)$, $x \in X$, is continuous and X is compact (see condition (iv)).

In the previous part of the paper, we have basically dealt with functionals $F(x(\cdot))$ representable in the form $F(x(\cdot)) = Ef(s, x(s))$, where $f(s, x)$ is the payoff function. In this section, we consider more general functionals, not necessarily representable in the above form. However, we here assume that the space S is finite, in contrast to Sections 2 and 3, where S was an arbitrary measurable space.

Lagrangian relaxation for the information constraint. We represent the problem $(\mathcal{P})$ in the equivalent form $(\mathcal{P}_0)$ and consider a Lagrange multiplier, $p(\cdot)$ $[p(\cdot) : S \to R^n]$, removing constraint (9). The existence of $p(\cdot)$ follows from Theorem 3.2. By definition, the function $p(\cdot)$ satisfies relation (11), which is equivalent to the following inequality

$$F(x(\cdot)) - E\ell(\cdot)x(\cdot) \leq F(\bar{x}) \qquad (x(\cdot) \in \mathcal{X}) \tag{118}$$

where

$$\ell(s) = Ep(\cdot) - p(s).$$

To establish the equivalence of (11) and (118), it suffices to use the identity

$$Ep(s)[x(s) - Ex(\cdot)] = E[p(s) - Ep(\cdot)]x(s).$$

We have

$$E\ell(\cdot) = 0, \tag{119}$$

so that $E\ell(\cdot)\bar{x} = 0$, and hence we conclude from (118) that

$$F(x(\cdot)) - E\ell(\cdot)x(\cdot) \leq F(\bar{x}) - E\ell(\cdot)\bar{x} = F(\bar{x}) \tag{120}$$

for all $x(\cdot) \in \mathcal{X}$.

Constraint (9) in the problem $(\mathcal{P}_0)$ does not allow the decision-maker to employ any decision functions except constants. This means that the decision-maker

cannot use information about s. In this section we will consider relaxations of constraint (9). These relaxations will correspond to certain (limited) possibilities of using information about s.

Decisions and their admissible corrections. To explain the idea of our approach, consider the model of realization of a risky project described in Sections 1 and 2. In that model, the project has to be realized during the period $[t_0, t_1]$. The manager takes a decision $x \in X$ at time t_0. Until the end of the project period, he does not receive any information about $s \in S$ which could be used for changing the initial plan. If the goal of the decision-maker is to maximize the functional $F(\cdot)$, then his optimal decisions are solutions to the problem $(\mathcal{P}_0)$.

Now suppose that there is a source of information which makes it possible to learn the exact value of $s \in S$ before the end of the project period $[t_0, t_1]$. Suppose the knowledge of s can be used for making a correction of the initial plan. This means that the initial decision $x \in X$ can be replaced by a decision function $x(s) = x + h(s)$, where $h(s)$ is the correction of the decision x at state s. For example, if the manager learns that the outcome of the project will be unfavorable, he may stop payments (or supply) in order to reduce future losses. On the other hand, if the outcome of the project turns out to be favorable, the manager may increase investments to get a greater reward. The earlier the decision-maker receives information about s, the more freedom for changing the initial decision he has, and so the larger is the class of admissible corrections $h(\cdot)$.

Based on the above considerations, we shall characterize a *source of information*, **I**, in terms of the class $\mathcal{H} = \mathcal{H}_{\mathbf{I}}$ of all *admissible corrections* $h(\cdot) : S \to R^n$ this source of information enables one to use. To the class $\mathcal{H}$ of admissible corrections, there corresponds the class

$$\mathcal{Y}(\mathcal{H}) = \{x(\cdot) \in \mathcal{X} : x(\cdot) = x + h(\cdot), x \in X, h(\cdot) \in \mathcal{H}\} \tag{121}$$

of admissible decision functions. According to (121), a function $x(\cdot) \in \mathcal{X}$ belongs to $\mathcal{Y}(\mathcal{H})$ if and only if $x(\cdot)$ is representable in the form $x(\cdot) = x + h(\cdot)$, where $x \in X$ is the initial decision and $h(\cdot) \in \mathcal{H}$ is the correction. The possibility of using a certain class $\mathcal{Y} = \mathcal{Y}(\mathcal{H})$ of *decision functions* (*strategies*) $x(\cdot) : S \to X$ (in addition to decisions $x \in X$ independent of s) will be regarded as a relaxation of the information constraint. Indeed, under the information constraint the only admissible correction is the zero correction.

The class of admissible corrections: some basic properties. Later on we will discuss the specification of the class $\mathcal{H} = \mathcal{H}_{\mathbf{I}}$ of admissible corrections in more detail. In this stage of exposition it is important to realize that without loss of generality one may restrict the attention to corrections $h(\cdot)$ belonging to the set $\mathcal{G}$, defined as

$$\mathcal{G} = \{h(\cdot) \in \mathcal{L} : \exists x \in X,\ x + h(s) \in X \text{ for all } s \in S\},$$

where $\mathcal{L}$ is the finite-dimensional linear space of all vector functions $h(\cdot) : S \to R^n$. Indeed, if $h(\cdot) \notin \mathcal{G}$ then $x + h(\cdot) \notin \mathcal{X}$ for all $x \in X$ so that such an $h(\cdot)$ gives no contribution to classes as $Y(\mathcal{H})$ defined in (121).

For $h(\cdot) \in \mathcal{G}$, the class

$$\mathcal{Y}(h(\cdot)) = \{x(\cdot) \in \mathcal{X} : x(\cdot) = x + h(\cdot),\ x \in X\}$$

is non-empty, and the maximum value of the objective functional $F(\cdot)$ which can be achieved by using the correction $h(\cdot) \in \mathcal{G}$ is defined as

$$\Phi(h(\cdot)) = \max_{x(\cdot)\in\mathcal{Y}(h(\cdot))} F(x(\cdot)).$$

Obviously, we have that $\mathcal{Y}(0)$ can be identified with X and $\Phi(0) = F(\bar{x})$. The larger is the value of the functional $\Phi(h(\cdot))$, the larger is the efficiency of the correction $h(\cdot)$.

In the next proposition some basic properties of $\mathcal{G}$, $\mathcal{Y}(h(\cdot))$ and $\Phi(h(\cdot))$ are proven.

PROPOSITION 7.1 *The set $\mathcal{G}$ is convex and compact; 0 is an interior point of $\mathcal{G}$. For each $h(\cdot) \in \mathcal{G}$, the set $\mathcal{Y}(h(\cdot))$ is non-empty, convex and compact. The functional $\Phi(h(\cdot))$, $h(\cdot) \in \mathcal{G}$, is (finite and) concave. If $\bar{x} \in \operatorname{int} X$ and $F(x(\cdot))$ is differentiable at $x(\cdot) = \bar{x}$, then $\Phi(h(\cdot))$ is differentiable at $h(\cdot) = 0$.*

Before proving this proposition, we introduce some notations. For two functions h_1 and h_2 in $\mathcal{L}$, we write

$$\langle h_1, h_2\rangle = Eh_1h_2.$$

We define

$$||h|| = \max_{s\in S} |h(s)|, h \in \mathcal{L}.$$

Differentiability of the functional $\Phi(h)$ at $h = 0$ means the existence of a function $\Phi'(0) = \Phi'(0)(s)$ in $\mathcal{L}$ such that

$$\Phi(h) - \Phi(0) = \langle \Phi'(0), h\rangle + \gamma(h), h \in \mathcal{G}, \tag{122}$$

where

$$\gamma(h)/||h|| \to 0 \tag{123}$$

as $||h|| \to 0$.

Proof of Proposition 7.1: The convexity and compactness of $\mathcal{G}$ and $\mathcal{Y}(h(\cdot))$ follow from the convexity and compactness of X. Let $\overset{\circ}{x} \in \operatorname{int} X$. Consider a number $\delta > 0$ such that $x \in X$ for all $x \in R^n$ with $|x - \overset{\circ}{x}| < \delta$. For any function $h(\cdot) \in \mathcal{L}$ with $||h(\cdot)|| < \delta$ we have $\overset{\circ}{x} + h(s) \in X$ $(s \in S)$, and so $h(\cdot) \in \mathcal{G}$. Thus $0 \in \operatorname{int} \mathcal{G}$.

Since $\mathcal{Y}(h(\cdot))$ is non-empty and compact for all $h(\cdot) \in \mathcal{G}$ it follows that the maximum in the definition of $\Phi(h(\cdot))$ is attained implying that $\Phi(h(\cdot))$ is finite for all $h(\cdot) \in \mathcal{G}$. To prove that $\Phi(\cdot)$ is concave, we fix $h_1(\cdot) \in \mathcal{G}$, $h_2(\cdot) \in \mathcal{G}$, $\theta \in [0, 1]$ and define $h(\cdot) = \theta h_1(\cdot) + (1 - \theta)h_2(\cdot)$. For $i = 1, 2$, let $x_i(\cdot)$ be a function in $\mathcal{Y}(h_i(\cdot))$ such that $\Phi(h_i(\cdot)) = F(x_i(\cdot))$. We have $x_i(\cdot) = x_i + h_i(\cdot) \in \mathcal{X}$ for some $x_i \in X$. Consequently,

$$x(\cdot) \equiv \theta x_1(\cdot) + (1 - \theta)x_2(\cdot) = x + h(\cdot) \in \mathcal{X},$$

where $x \equiv \theta x_1 + (1-\theta)x_2 \in X$. Thus, $x(\cdot) \in \mathcal{Y}(h(\cdot))$, and we can write

$$\Phi(h(\cdot)) \geq F(x(\cdot)) \geq \theta\, F(x_1(\cdot)) + (1-\theta)F(x_2(\cdot)) = \theta\, \Phi(h_1(\cdot)) + (1-\theta)\Phi(h_2(\cdot)),$$

which proves the desired statement.

Denote by $\mathcal{L}^*$ the dual of $\mathcal{L}$. Consider the set

$$\partial\Phi(0) = \{m \in \mathcal{L}^* : \Phi(h(\cdot)) - \Phi(0) \leq \langle m, h\rangle,\ h(\cdot) \in \mathcal{G}\}$$

of support functionals of the functional $\Phi(h(\cdot))$ at $h(\cdot) = 0$. Since $0 \in \operatorname{int} \mathcal{G}$ and $\Phi(h(\cdot))$ is concave, we have $\partial\Phi(0) \neq \emptyset$. A concave functional is differentiable at an interior point of its domain if and only if it has a unique support functional at this point (see Rockafellar [20], Theorem 25.1). Thus, it suffices to prove that $\partial\Phi(0)$ is a singleton. Let $m \in \partial\Phi(0)$. Consider any function $x(\cdot) \in \mathcal{X}$ and define $h(\cdot) = x(\cdot) - \bar{x}$. We have $h(\cdot) \in \mathcal{G}$ and $x(\cdot) \in \mathcal{Y}(h(\cdot))$. Hence

$$\langle m, x(\cdot) - \bar{x}\rangle = \langle m, h(\cdot)\rangle \geq \Phi(h(\cdot)) - \Phi(0) \geq F(x(\cdot)) - F(\bar{x}).$$

Consequently, $m \in \partial F(\bar{x})$, and so $m = F'(\bar{x})$, since the functional $F(x(\cdot))$ is differentiable at the interior point $x(\cdot) = \bar{x}$ of the set $\mathcal{X}$. □

We can see from the above proof of Proposition 7.1 that $\Phi'(0) = F'(\bar{x})$, provided $\bar{x} \in \operatorname{int} X$ and $F(x(\cdot))$ is differentiable at $x(\cdot) = \bar{x}$. Therefore, if $F(x(\cdot)) = Ef(s, x(s))$, where $f(s,x)$ is differentiable at $x = \bar{x} \in \operatorname{int} X$ for each s, then we have $\Phi'(0)(s) = F'(\bar{x})(s) = f'_x(s, \bar{x})$.

Flexibility of decision functions. Let **I** be a source of information characterized by a class $\mathcal{H} = \mathcal{H}_{\mathbf{I}} \subseteq \mathcal{G}$ of admissible corrections. Consider the corresponding class $\mathcal{Y} = \mathcal{Y}(\mathcal{H}_{\mathbf{I}})$ of admissible decision functions (strategies) $x(\cdot)$ in $\mathcal{X}$. We wish to define a quantitative characteristic of efficiency of the source of information. To this end we will use a quantitative measure for the notion of *flexibility* of admissible strategies. Typically, we will consider sources of information that allow us to use a strategy precisely if its flexibility does not exceed a prescribed upper bound. Therefore, the efficiency of such a source of information increases with the bound on the flexibility.

How to specify the flexibility of a strategy? Intuitively, the flexibility of a strategy $x(\cdot)$ is a measure of the dependence of $x(s)$ on s. If $x(s) =$ constant, then its flexibility should be equal to zero. The larger the spread of values of $x(s)$, $s \in S$, is, the larger should the flexibility of $x(\cdot)$ be. We employ a functional $\mathcal{V}(x(\cdot))$ (which will be defined axiomatically below) for measuring the "degree of utilization" of information about s in the strategy $x(s)$.

If $x(s)$ does not depend on s (i.e., $\mathcal{V}(x(\cdot)) = 0$), then no information about s is used in $x(\cdot)$. If there is a significant distinction between the values of $x(s)$ for different states s (say, for favorable and unfavorable outcomes), then this means that the knowledge of s plays a significant role in the process of decision making.

Axioms for a measure of flexibility. In different concrete problems, it may be convenient to consider different functionals $\mathcal{V}(x(\cdot))$. Therefore we shall not assume that $\mathcal{V}(x(\cdot))$ is given in one or another specific form. Rather, we shall present a list

of axioms describing those properties of $\mathcal{V}(x(\cdot))$ which are needed for our purposes. These axioms are as follows.

($\mathcal{V}$.0) The functional $\mathcal{V}(x(\cdot))$ is defined for any function $x(\cdot) \in \mathcal{L}$ and takes non-negative (finite) values.

($\mathcal{V}$.1) For any $y \in R^n$ and $x(\cdot) \in \mathcal{L}$, we have $\mathcal{V}(y + x(\cdot)) = \mathcal{V}(x(\cdot))$.

($\mathcal{V}$.2) If $x(\cdot) \neq$ constant, then $\mathcal{V}(x(\cdot)) > 0$.

($\mathcal{V}$.3) We have $\mathcal{V}(\lambda x(\cdot)) = \lambda \mathcal{V}(x(\cdot))$ for any scalar $\lambda \geq 0$ and $x(\cdot) \in \mathcal{L}$.

($\mathcal{V}$.4) The functional $\mathcal{V}(x(\cdot))$, $x(\cdot) \in \mathcal{L}$, is continuous (with respect to the conventional topology in the finite-dimensional space $\mathcal{L}$).

Notice that these axioms imply that $\mathcal{V}(x(\cdot)) = 0$ if $x(\cdot) =$ constant. Indeed, by taking $\lambda = 0$ in ($\mathcal{V}$.3) we get $\mathcal{V}(0) = 0$; and application of ($\mathcal{V}$.1) gives the desired result.

An example of a functional $\mathcal{V}(x(\cdot))$ with properties ($\mathcal{V}$.0) – ($\mathcal{V}$.4) is as follows:

$$\mathcal{V}(x(\cdot)) = \mathcal{N}(x(\cdot) - Ex(\cdot)), \tag{124}$$

where $\mathcal{N}$ is any norm in the space $\mathcal{L}$. Observe that if $\mathcal{V}(x(\cdot))$ is representable in the form (125), then the set $\{x(\cdot) \in \mathcal{L} : \mathcal{V}(x(\cdot)) \leq 1\}$ is convex and symmetric with respect to zero. This is not necessarily so for general functionals satisfying ($\mathcal{V}$.0) – ($\mathcal{V}$.4).

It follows from ($\mathcal{V}$.2) – ($\mathcal{V}$.4) that there exist $\theta > 0$ and $\Theta > 0$ satisfying

$$\theta ||x(\cdot)|| \leq \mathcal{V}(x(\cdot)) \leq \Theta ||x(\cdot)|| \tag{125}$$

for all $x(\cdot) \in \mathcal{L}$ with $Ex(\cdot) = 0$. Indeed, let θ and Θ be the minimum and the maximum values of $\mathcal{V}(x(\cdot))$ on $\{x(\cdot) \in \mathcal{L} : ||x(\cdot)|| = 1,\ Ex(\cdot) = 0\}$. These values are attained by virtue of ($\mathcal{V}$.4). Thus, $\theta = \mathcal{V}(x_0(\cdot))$ for some $x_0 \in \mathcal{L}$, where $||x_0(\cdot)|| = 1$ and $Ex_0(\cdot) = 0$ so that x_0 is not a constant function. In view of ($\mathcal{V}$.2), we have $\theta > 0$. By using ($\mathcal{V}$.3), we establish the first inequality in (125). The second one is derived similarly.

We fix a functional $\mathcal{V}(x(\cdot))$, $x(\cdot) \in \mathcal{L}$, satisfying ($\mathcal{V}$.0) – ($\mathcal{V}$.4). This functional will serve as a measure of flexibility of admissible strategies or corrections. For a class $\mathcal{Z}$ of functions $x(\cdot) \in \mathcal{L}$, we denote the maximum flexibility of the functions in $\mathcal{Z}$ as

$$\mathcal{V}(\mathcal{Z}) = \sup_{x(\cdot) \in \mathcal{Z}} \mathcal{V}(x(\cdot)).$$

The next proposition shows that the maximum flexibility of a class of admissible corrections is equal to the maximum flexibility of the corresponding class of strategies.

PROPOSITION 7.2 *Let $\mathcal{H} \subseteq \mathcal{G}$ be a class of admissible corrections and $\mathcal{Y}(\mathcal{H})$ the corresponding class of admissible decision functions. Then we have*

$$\mathcal{V}(\mathcal{H}) = \mathcal{V}(\mathcal{Y}(\mathcal{H})). \tag{126}$$

PROOF: If $x(\cdot) \in \mathcal{Y}(\mathcal{H})$, then $x(\cdot) = x + h(\cdot)$ for some $x \in X$ and $h(\cdot) \in \mathcal{H}$. By virtue of $(\mathcal{V}.1)$, we have $\mathcal{V}(x(\cdot)) = \mathcal{V}(h(\cdot)) \leq \mathcal{V}(\mathcal{H})$. Thus $\mathcal{V}(\mathcal{Y}(\mathcal{H})) \leq \mathcal{V}(\mathcal{H})$. On the other hand, if $h(\cdot) \in \mathcal{H} \subseteq \mathcal{G}$, then $x + h(\cdot) \in \mathcal{Y}(\mathcal{H})$ for some $x \in X$. Consequently, $\mathcal{V}(h(\cdot)) = \mathcal{V}(x + h(\cdot)) \leq \mathcal{V}(\mathcal{Y}(\mathcal{H}))$, and so $\mathcal{V}(\mathcal{H}) \leq \mathcal{V}(\mathcal{Y}(\mathcal{H}))$, which completes the proof. □

The class of corrections of flexibility $\leq r$. Consider the class

$$\mathcal{H}_r = \{h(\cdot) \in \mathcal{G} : \mathcal{V}(h(\cdot)) \leq r\} \ (r \geq 0) \tag{127}$$

of all admissible corrections whose flexibility is not greater than r. Let

$$\mathcal{Y}_r = \{x(\cdot) \in \mathcal{X} : \mathcal{V}(x(\cdot)) \leq r\} \ (r \geq 0) \tag{128}$$

be the class of all strategies of flexibility $\leq r$. Define

$$\mathcal{H}_r^0 = \mathcal{H}_r \cap \{h(\cdot) : Eh(\cdot) = 0\}. \tag{129}$$

Note that the sets $\mathcal{H}_r$, $\mathcal{Y}_r$ and $\mathcal{H}_r^0$ are compact, since $\mathcal{G}$ and $\mathcal{X}$ are compact and $\mathcal{V}(\cdot)$ is continuous.

PROPOSITION 7.3 *We have*

$$\mathcal{Y}_r = \mathcal{Y}(\mathcal{H}_r) = \mathcal{Y}(\mathcal{H}_r^0) \ (r \geq 0). \tag{130}$$

Proof: Clearly $\mathcal{Y}(\mathcal{H}_r^0) \subseteq \mathcal{Y}(\mathcal{H}_r)$. Let $x(\cdot) \in \mathcal{Y}(\mathcal{H}_r)$. Then $x(\cdot) = x + h(\cdot) \in \mathcal{X}$, where $\mathcal{V}(h(\cdot)) \leq r$. Therefore $\mathcal{V}(x(\cdot)) = \mathcal{V}(h(\cdot)) \leq r$, and so $x(\cdot) \in \mathcal{Y}_r$. Thus $\mathcal{Y}(\mathcal{H}_r^0) \subseteq \mathcal{Y}(\mathcal{H}_r) \subseteq \mathcal{Y}_r$.

Let $x(\cdot) \in \mathcal{Y}_r$. Then $x(s) \in X$ for all s, and hence $\hat{x} \equiv Ex(\cdot) \in X$ by virtue of the convexity of X. Define $h(\cdot) = x(\cdot) - \hat{x}$. Then we have

$$x(\cdot) = \hat{x} + h(\cdot) \in \mathcal{X}, \ \hat{x} \in X,$$

and $Eh(\cdot) = 0$. Furthermore, $\mathcal{V}(h(\cdot)) = \mathcal{V}(x(\cdot)) \leq r$. Consequently, $x(\cdot) \in \mathcal{Y}(\mathcal{H}_r^0)$. □

The conjugate to the functional $\mathcal{V}$. For any $g(\cdot) \in \mathcal{L}$, we define

$$\mathcal{V}^*(g(\cdot)) = \max_{h(\cdot) \in \mathcal{L}_1^0} Eg(\cdot)h(\cdot), \tag{131}$$

where

$$\mathcal{L}_1^0 = \{h(\cdot) \in \mathcal{L} : Eh(\cdot) = 0, \ \mathcal{V}(h(\cdot)) \leq 1\}. \tag{132}$$

It follows from (125) that the set $\mathcal{L}_1^0$ is compact. Therefore the maximum in (131) is attained. The functional $\mathcal{V}^*(\cdot)$ is termed *conjugate* to $\mathcal{V}(\cdot)$. It can be shown that $\mathcal{V}^*(\cdot)$ is concave and satisfies $(\mathcal{V}.0) - (\mathcal{V}.4)$. The proof is left to the reader.

Let $\mathbf{I}_r$ $(r \geq 0)$ be a source of information which enables one to use all the corrections in the class $\mathcal{H}_r$. By virtue of (130), this source of information makes it possible to use all the strategies in the class $\mathcal{Y}_r$, i.e., all the admissible strategies of flexibility $\leq r$. Let

$$\varphi(r) = \max_{x(\cdot)\in\mathcal{Y}_r} F(x(\cdot)) \ (r \geq 0) \tag{133}$$

be the maximum value of the objective functional on the class $\mathcal{Y}_r$.

The key result. The main result of this section is as follows.

THEOREM 7.1 *Let $\ell(\cdot)$ be a vector function satisfying conditions (118) and (119). Then we have*

$$\varphi(r) - \varphi(0) \leq \mathcal{V}^*(\ell(\cdot)) \cdot r \tag{134}$$

for any $r \geq 0$. Moreover, if the following condition holds

(Φ) *the functional $\Phi(h(\cdot))$ is differentiable at $h(\cdot)) = 0$,*

then

$$\varphi(r) - \varphi(0) = \mathcal{V}^*(\ell(\cdot)) \cdot r + o(r) \ (r \geq 0). \tag{135}$$

As usual, $o(r)$ stands for a function of r which tends to zero faster than r as $r \to 0$. By virtue of Proposition 7.1, condition (Φ) holds if the functional $F(x(\cdot))$ is differentiable at $x(\cdot) = \bar{x}$.

Before proving the above theorem, we discuss its economic meaning.

The shadow price of information. Suppose the decision-maker who plans to implement the risky project can get access to information about the random parameter $s \in S$ at a certain price. Assume the decision-maker can "buy" one or another source of information in the class $\{\mathbf{I}_r,\ r \geq 0\}$. Let the cost of a source of information be proportional to its efficiency. Then the cost of $\mathbf{I}_r$ is λr, where λ is a constant. This constant may be regarded as a price for a unit of efficiency of a source of information. We say that λ is a *shadow price of information* if the following conditions hold for $r \geq 0$:

$$\varphi(r) - \varphi(0) \ \leq \ \lambda r; \tag{136}$$

$$\varphi(r) - \varphi(0) \ = \ \lambda r + o(r). \tag{137}$$

The first of these requirements states that the increase of the objective functional yielded by the source $\mathbf{I}_r$ is not greater than the cost λr of this source. By virtue of the second requirement, the increment of the functional $\varphi(\cdot)$ is approximately equal to λr for small r. Relation (137) can be rewritten in the form

$$\frac{\varphi(r) - \varphi(0)}{r} \to \lambda$$

as $r \downarrow 0$. This means that λ is equal to the marginal increase of the objective functional under a marginal relaxation of the information constraint. Thus, relations (136) and (137) express fundamental properties of shadow prices which are characteristic for this concept (e.g. see Birchenhall and Grout [5], Chapter 9).

Set $\lambda = \mathcal{V}^*(\ell(\cdot))$. Then, under assumption (Φ), (134) and (135) imply (136) and (137). Therefore $\mathcal{V}^*(\ell(\cdot))$ may be interpreted as a shadow price of information.

Let us return for the moment to the robust insurance mechanisms considered in the previous part of the paper. In the insurance scheme defined in Section 2, the vector function $\ell(s) = Ep(\cdot) - p(s)$ represents the difference between the premium price vector $\tilde{p} = Ep(\cdot)$ and the compensation price vector $p(s)$. Thus, Theorem 7.1 establishes a direct connection between the shadow prices on information and the robust insurance mechanisms.

Proof of Theorem 7.1: It follows from (118) that

$$F(x(\cdot)) \leq F(\bar{x}) + E\ell(\cdot)x(\cdot) \tag{138}$$

for all $x(\cdot) \in \mathcal{X}$. Consequently,

$$\phi(r) \leq F(\bar{x}) + \max_{x(\cdot)\in\mathcal{Y}_r} E\ell(\cdot)x(\cdot) = \varphi(0) + \max_{x(\cdot)\in\mathcal{Y}_r} E\ell(\cdot)x(\cdot). \tag{139}$$

By virtue of Proposition 7.3, any function $x(\cdot) \in \mathcal{Y}_r$ is representable in the form $x(\cdot) = x + h(\cdot)$, where $x \in X$ and $h(\cdot) \in \mathcal{H}_r^0$. Since $E\ell(\cdot) = 0$ (see (119)), we have $E\ell(\cdot)x(\cdot) = E\ell(\cdot)h(\cdot)$. Therefore

$$\max_{x(\cdot)\in\mathcal{Y}_r} E\ell(\cdot)x(\cdot) \leq \max_{h(\cdot)\in\mathcal{L}_r^0} E\ell(\cdot)h(\cdot) = r \cdot \max_{h(\cdot)\in\mathcal{L}_1^0} E\ell(\cdot)h(\cdot) = r \cdot \mathcal{V}^*(\ell(\cdot)), \tag{140}$$

where

$$\mathcal{L}_r^0 = \{h(\cdot) \in \mathcal{L} : \ Eh(\cdot) = 0, \ \mathcal{V}(h(\cdot)) \leq r\}.$$

By combining (139) and (140), we obtain (134).

In order to prove (135) we assume that condition (Φ) holds. For any $h(\cdot) \in \mathcal{G}$, we have because of (118)

$$\Phi(h(\cdot)) \leq F(\bar{x}) + \max_{x(\cdot)\in\mathcal{Y}(h(\cdot))} E\ell(\cdot)x(\cdot) = \Phi(0) + E\ell(\cdot)h(\cdot),$$

since $E\ell(\cdot)x(\cdot) = E\ell(\cdot)h(\cdot)$ for each $x(\cdot) \in \mathcal{Y}(h(\cdot))$ because of (119). Consequently, $\ell(\cdot) \in \partial\Phi(0)$, and since (Φ) holds, actually $\ell(\cdot) = \Phi'(0)$. Therefore

$$\Phi(h(\cdot)) = \Phi(0) + E\ell(\cdot)h(\cdot) + \gamma(h(\cdot)), \tag{141}$$

where

$$\gamma(h(\cdot))/||h(\cdot)|| \to 0 \tag{142}$$

as $||h(\cdot)|| \to 0$.

By using Proposition 7.3, we write

$$\varphi(r) = \max_{x(\cdot)\in\mathcal{Y}(\mathcal{H}_r^0)} F(x(\cdot)) = \max_{h(\cdot)\in\mathcal{H}_r^0} \max_{x(\cdot)\in\mathcal{Y}(h(\cdot))} F(x(\cdot)) = \max_{h(\cdot)\in\mathcal{H}_r^0} \Phi(h(\cdot)). \tag{143}$$

Observe that $\mathcal{H}_r^0 = \mathcal{L}_r^0$ for all $r > 0$ small enough. Indeed,

$$\mathcal{H}_r^0 = \mathcal{G} \cap \mathcal{L}_r^0 \subseteq \mathcal{L}_r^0$$

for each $r \geq 0$. Since $0 \in \operatorname{int} \mathcal{G}$, there exists an $\varepsilon > 0$ such that $\{h(\cdot) \in \mathcal{G} : ||h(\cdot)|| < \varepsilon\} \subseteq \mathcal{G}$. Then, by virtue of (125), we have $\mathcal{L}_r^0 \subseteq \mathcal{G}$ if $r \in [0, r_0)$, where $r_0 = \theta\varepsilon$. Thus, $\mathcal{H}_r^0 = \mathcal{L}_r^0$ for all $r \in [0, r_0)$. By using this, combined with (141) and (143), we find

$$\varphi(r) - \varphi(0) = \max_{h(\cdot) \in \mathcal{L}_r^0} [E\ell(\cdot)h(\cdot) + \gamma(h(\cdot))] \ (0 \leq r < r_0).$$

Consequently, using $\mathcal{L}_r^0 = r \cdot \mathcal{L}_1^0$ and (131), we get

$$\varphi(r) - \varphi(0) = r \cdot \mathcal{V}^*(\ell(\cdot)) + \zeta(r),$$

where

$$|\zeta(r)| \leq \sup_{h(\cdot) \in \mathcal{L}_r^0} |\gamma(h(\cdot))|. \tag{144}$$

We conclude from (144), (142) and (125) that $|\zeta(r)|/r \to 0$ as $r \to 0$. □

Comparison with the conventional theory. Let us compare our approach to the analysis of decision functions and information with a more traditional theory. According to the standard methodology, limited possibilities of using information about $s \in S$ are described as follows. It is assumed that the decision-maker can use only those strategies which are measurable with respect to a given partition $\mathcal{S}$ of the space S. The partition $\mathcal{S}$ is specified by a collection of sets $S_1, \ldots, S_k \subseteq S$ such that $S = S_1 \cup \ldots \cup S_k$ and $S_i \cap S_j = \emptyset$ for $i \neq j$. A function $x(\cdot) : S \to R^n$ is measurable with respect to $\mathcal{S}$ if $x(s') = x(s'')$ for s', $s'' \in S_i$, $i = 1, \ldots, k$ (i.e., $x(\cdot)$ is constant on each of the sets S_i). The decision-maker does not know the true value of s exactly, but he knows to which of the sets S_i the value of s belongs. Therefore the admissible decision functions are measurable with respect to $\mathcal{S}$. For a general – not necessarily finite – space S, similar concepts are usually defined in terms of σ-algebras on S (e.g. see Radner [19] and Allen [1], where the subject is discussed in connection with economic theory).

In contrast to the conventional methodology, we assume that the decision-maker may know the exact value of s, but he cannot use this knowledge as freely as he might wish (for example, because the message about s comes too late). The decision-maker is restricted to a certain class of decision functions with a given maximal level of flexibility. This level of flexibility is regarded as a fundamental characteristic of the source of information. Such an approach enables us to formalize the idea of "useful" or "effective" information about s which is actually used in the process of decision making.

Note that the above-mentioned standard approach cannot directly be applied to the analysis of our problem, in particular, for the following reason. A key role in our considerations is played by the idea of a marginal relaxation of the information constraint. Therefore we have to be able to consider arbitrarily small and infinitely

divisible portions of information. If S is finite (which has been assumed in this section), then there is only a finite number of partitions of S. Consequently, the "amount of information" can vary by discrete portions only. Furthermore, if S consists of two points (success and failure), then there are only two partitions of S: the trivial one $\{S, \emptyset\}$, and the partition of S into points. In other words, there is either no information or complete information about s; the conventional approach does not enable us to consider any intermediate cases.

It would be of interest to extend the results of this section to more general models. In particular, one can try to analyze the concept of shadow prices on information in the framework of dynamic (multi-stage) problems of decision making under uncertainty. Another interesting field for possible applications and generalizations of the methods developed is the theory of teams (Marschak and Radner [17]).

Acknowledgments

We wish to thank B. Allen, K.J. Arrow, M.J. Goovaerts, K. Schmidt, C. Schweigman, U. Schweitzer, D. Sondermann and R.J.-B. Wets for helpful comments. We are especially grateful to J.H. Drèze for numerous stimulating discussions and valuable remarks. Ideas and suggestions of J.H. Drèze have been used in this work. Financial support from Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at the University of Bonn, from MacArthur Foundation Grant 94-26979A-FSU, and from INTAS Grant 97-1050 is gratefully acknowledged.

Appendix

In the Appendix we present proofs of Theorem 3.2 and Lemmas 6.1 and 6.2.

The proof of theorem 3.2 relies on a general fact from functional analysis. Let D_1 and D_2 be Banach spaces, Z a convex subset of D_1, $F : Z \to (-\infty, +\infty)$ a concave functional and $B : D_1 \to D_2$ a continuous linear operator. Assume that the functional $F(z)$ attains its maximum on the set $\{z \in Z : Bz = 0\}$ at some point $\bar{z}$.

LEMMA A.1 *Let the following conditions hold.*

(A) The set $\{z \in Z : Bz = 0\}$ contains a point $\overset{\circ}{z}$ belonging to the interior of the set Z.

(B) The image of D_1 under the mapping $B : D_1 \to D_2$ is a closed subspace of D_2.

(C) The functional $F(\cdot)$ is bounded below on some open subset of Z.

Then there exists a continuous linear functional π on the space D_2 such that

$$F(z) + \langle \pi,\ Bz \rangle \leq F(\bar{z}) \tag{A.1}$$

for each $z \in Z$.

This result is an infinite-dimensional version of the Kuhn-Tucker theorem for concave optimization problems with linear equality constraints. A proof of Lemma

A.1 is presented in [10], Proposition FA1, see also [16]. Note that condition (C) holds if $F(\cdot)$ is continuous on int Z.

Proof of Theorem 3.2: Denote by L_∞ the space of essentially bounded (with respect to the measure μ) vector functions $x(s)$, $s \in S$, with values in R^n. We wish to apply Lemma A.1 to the Banach spaces $D_1 = D_2 = L_\infty$, the operator $Bx(\cdot) = x(\cdot) - Ex(\cdot)$ for $x(\cdot) \in D_1$, the set $Z = \mathcal{X}$ and the functional $F(x(\cdot))$, $x(\cdot) \in \mathcal{X}$, where $\mathcal{X}$ and F are described in Theorem 3.2. The functional F is continuous with respect to a.s. convergence, and hence F is continuous in the norm $||x(\cdot)||_\infty = \text{ess sup } |x(s)|$, $x(\cdot) \in D_1$. Let $\overset{\circ}{x}$ be an element of int X. Since X is convex and $\overset{\circ}{x} \in$ int X, the set $\mathcal{X}$ is convex and the constant vector function $\overset{\circ}{x}$ belongs to the interior of $\mathcal{X}$. Furthermore, we have $B\overset{\circ}{x} = \overset{\circ}{x} - E\overset{\circ}{x} = 0$. The operator B is continuous in the norm $||\cdot||_\infty$ and maps D_1 onto the linear space $D_3 = \{y(\cdot) \in D_2 : Ey(\cdot) = 0\}$, which is closed in D_2. The functional $F(x(\cdot))$ attains its maximum on the set $\{x(\cdot) \in \mathcal{X} : Bx(\cdot) = 0\}$, since this set is compact in L_∞ (its elements can be identified with vectors $x \in X$) and since $F(\cdot)$ is continuous in the norm $||\cdot||_\infty$. Thus, all the conditions of Lemma A.1 are fulfilled. By using this lemma, we conclude that there is a functional $\pi \in L_\infty^*$ satisfying

$$F(x(\cdot)) + \langle \pi, x(\cdot) - Ex(\cdot)\rangle \leq F(\bar{x}) \ (x(\cdot) \in X) \tag{A.2}$$

where $\bar{x}$ is a solution to the problem $(\mathcal{P}_0)$.

By virtue of the Yosida-Hewitt theorem [26], any functional $\pi \in L_\infty^*$ can be decomposed into the sum $\pi = \pi^a + \pi^s$ of two functionals $\pi^a \in L_\infty^*$ and $\pi^s \in L_\infty^*$ possessing the following properties:

(π.1) There exists a vector function $p(s)$, $s \in S$, such that $E|p(\cdot)| < \infty$ and

$$\langle \pi^a, x(\cdot)\rangle = Ep(s)x(s), \ x(\cdot) \in L_\infty.$$

(π.2) There exists a sequence of measurable sets $S \supseteq \Gamma_1 \supseteq \Gamma_2 \supseteq \ldots$ such that $\mu(\Gamma_k) \to 0$ $(k \to \infty)$ and $\langle \pi^s, x(\cdot)\rangle = \langle \pi^s, \mathbf{1}_{\Gamma_k} x(\cdot)\rangle$ for all $k = 1, 2, \ldots$ and $x(\cdot) \in L_\infty$ (we denote by $\mathbf{1}_\Gamma$ the indicator function of the set Γ).

Fix any $x(\cdot) \in \mathcal{X}$ and define

$$x_k(s) = (1 - \mathbf{1}_{\Gamma_k}(s))x(s) + \mathbf{1}_{\Gamma_k}(s) \cdot Ex(\cdot), \ k = 1, 2, \ldots$$

Since $x(s) \in X$ a.s. and X is a closed convex set, we have $Ex(\cdot) \in X$ (see, e.g., [2], Appendix II). Consequently, $x_k(\cdot) \in \mathcal{X}$. By using (A.2), we find

$$F(x_k(\cdot)) + \langle \pi^a, x_k(\cdot) - Ex_k(\cdot)\rangle + \langle \pi^s, x_k(\cdot) - Ex_k(\cdot)\rangle \leq F(\bar{x}). \tag{A.3}$$

Here, $F(x_k(\cdot)) \to F(x(\cdot))$ and

$$\langle \pi^a, x_k(\cdot) - Ex_k(\cdot)\rangle \to \langle \pi^a, x(\cdot) - Ex(\cdot)\rangle = Ep(s)(x(s) - Ex(\cdot))$$

as $k \to \infty$, since $x_k(\cdot) \to x(\cdot)$ a.s. Furthermore,

$$\begin{aligned} \langle \pi^s, x_k(\cdot) - Ex_k(\cdot)\rangle &= \langle \pi^s, \mathbf{1}_{\Gamma_k}(x_k(\cdot) - Ex_k(\cdot))\rangle = \\ \langle \pi^s, \mathbf{1}_{\Gamma_k}(Ex(\cdot) - Ex_k(\cdot))\rangle &= \langle \pi^s, Ex - Ex_k(\cdot)\rangle \to 0 \end{aligned}$$

$(k \to \infty)$. Thus, by passing to the limit in (A.3), we obtain

$$F(x(\cdot)) + Ep(s)(x(s) - Ex(\cdot)) \leq F(\bar{x}),$$

which proves the theorem. □

For a probability measure γ on a Euclidean space R^d, we denote by $\operatorname{supp}\gamma$ the support of γ, i.e., the smallest closed set $V \subseteq R^d$ with $\gamma(V) = 1$.

Proof of Lemma 6.1: Let μ_0 denote the image of the measure μ under the mapping $\zeta : S \to R^d$. Denote by m_0 the image of the measure m under the mapping $(\zeta \otimes \zeta) : S \times S \to R^d \times R^d$, transforming (s_1, s_2) into $(\zeta(s_1), \zeta(s_2))$. We have $m_0 = \mu_0 \times \mu_0$, and so $\operatorname{supp} m_0 = \operatorname{supp}\mu_0 \times \operatorname{supp}\mu_0$. Since W is closed and $m_0(W) = 1$ (see (105)), we obtain that $\operatorname{supp} m_0 \subseteq W$. Consider the set $S^* = \{s \in S : \zeta(s) \in \operatorname{supp}\mu_0\}$, fix some point $s^* \in S^*$, and define $\zeta'(s) = \zeta(s)$ if $s \in S^*$ and $\zeta'(s) = \zeta(s^*)$ if $s \in S \setminus S^*$. By the definition of S^*, we have $\mu(S^*) = 1$, which implies (106). From the properties of $\operatorname{supp} m_0$ mentioned above, it follows that (107) holds for all $(s_1, s_2) \in S \times S$. □

Proof of Lemma 6.2: Suppose $e(r) \neq r$ for any $r \in R^1$. Then either $e(r) > r$ for all r, or $e(r) < r$ for all r. In both cases, the equality in (110) cannot hold. Consequently, there is a number r^* satisfying $e(r^*) = r^*$. From this, (108) and (109), it follows that $e(r) \geq r$ if $r \leq r^*$ and $e(r) \leq r$ if $r \geq r^*$. Denote by $\kappa_1(dr)$ the image of the measure $\kappa(dr)$ under the mapping $e : R^1 \to R^1$. Consider the distribution functions $D(r) = \kappa((-\infty, r])$ and $D_1(r) = \kappa_1((-\infty, r])$ of the measures $\kappa(dr)$ and $\kappa_1(dr)$. We have $D_1(r) = \kappa(\{r' : e(r') \leq r\})$. By using the relation $(r - e(r))(r - r^*) \geq 0$ and the properties (108), (109) of the function $e(\cdot)$, it is not hard to prove the following: $\{r' : e(r') \leq r\} \subseteq (-\infty, r]$ if $r < r^*$ and $\{r' : e(r') \leq r\} \supseteq (-\infty, r]$ when $r \geq r^*$. From this we conclude

$$D_1(r) \leq D(r), \ r < r^*; \quad D_1(r) \geq D(r), \ r \geq r^*. \tag{A.4}$$

It is known that properties (A.4) and (110) imply (111) (see, e.g., Hanoch and Levy [12], Theorem 3). □

References

1. B. Allen. Neighboring information and distribution of agents' characteristics under uncertainty. *Journal of Mathematical Economics*, 12: 63–101, 1983.
2. V.I. Arkin and I.V. Evstigneev. *Stochastic Models of Control and Economic Dynamics.* Academic Press, London, 1987.
3. K.J. Arrow. *Essays in the Theory of Risk-Bearing.* North Holland, Amsterdam, 1971.
4. K. Back and S.R. Pliska. The shadow price of information in continuous time decision problems. *Stochastics*, 22: 151–186, 1987.
5. C. Birchenhall and P. Grout. *Mathematics for Modern Economics.* Philip Allan Publ., Oxford, 1984.
6. D. Blackwell. Comparison of experiments. In: *Proc. Second Berkeley Symp. on Mathematical Statistics and Probability (1950).* University of California Press, Berkeley, 93–102, 1951.
7. K.H. Borch. *Economics of Insurance.* North Holland, Amsterdam, 1990.
8. J.H. Drèze. *Essays on Economic Decisions under Uncertainty.* Cambridge Univ. Press, Cambridge, 1987.

9. N. Dunford and J.T. Schwartz. *Linear Operators, Part 1, General Theory.* Interscience Publ., New York, 1957.
10. I.V. Evstigneev. Lagrange multipliers for the problems of stochastic programming. In: *Warsaw Fall Seminars in Mathematical Economics (1975)*, M.W. Loś, J. Loś and A. Wieczorek, eds., Lect. Notes in Econ. and Math. Syst., No. 133. Springer-Verlag, Berlin, 34–48, 1976.
11. S. Grossman and O. Hart. An analysis of the principal–agent problem. *Econometrica*, 51: 7–45, 1983.
12. G. Hanoch and H. Levy. The efficiency analysis of choices involving risk. *Review of Economic Studies*, 36: 335–346, 1969.
13. O. Hart and B. Holmström. The theory of contracts. In: *Advances in Economic Theory: Fifth World Congress*, T. Bewley, ed., Cambridge University Press, Cambridge, UK, 71–156, 1987.
14. A.D. Ioffe and V.M. Tihomirov. On the minimization of integral functionals. *Functional Analysis*, 3: 61–70, 1969.
15. A.D. Ioffe and V.M. Tihomirov. *Theory of Extremal Problems.* North Holland, New York, 1979.
16. V.L. Levin. *Convex Analysis in Spaces of Measurable Functions.* Nauka, Moscow, 1985. (In Russian.)
17. J. Marschak and R. Radner. *Economic Theory of Teams.* Yale University Press, New Haven and London, 1972.
18. P. Milgrom and J. Roberts. *Economics, Organization and Management.* Prentice Hall, Englewood Cliffs, N.J., 1992.
19. R. Radner. Equilibrium under uncertainty. In: *Handbook of Mathematical Economics*, K.J. Arrow and M.D. Intrilligator, eds., North Holland, Amsterdam, 923–1006, 1982.
20. R.T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, 1970.
21. R.T. Rockafellar and R.J.-B. Wets. Non-anticipativity and L^1-martingales in stochastic optimization problems. *Math. Programming Study*, 6: 170–187, 1976.
22. R.T. Rockafellar and R.J.-B. Wets. On the interchange of subdifferentiation and conditional expectation for convex functions. *Stochastics*, 7: 173–182, 1982.
23. M. Rothschild and J.E. Stiglitz. Increasing risk: I. A definition. *Journal of Economic Theory*, 2: 225–243, 1970.
24. D. Sondermann. Reinsurance in arbitrage-free markets. *Insurance: Mathematics and Economics*, 10: 191–202, 1991.
25. A.M. Yaglom and I.M. Yaglom. *Probability and Information.* Reidel, Dordrecht, 1983.
26. K. Yosida and E. Hewitt. Finitely additive measures. *Trans. Amer. Math. Soc.*, 72: 46–66, 1952.