

# PARETO-OPTIMAL RISK EXCHANGES AND RELATED DECISION PROBLEMS

HANS U. GERBER

## I. SUMMARY

In various branches of applied mathematics the problem arises of making decisions to reconcile conflicting criteria. One example is the classical statistical problem, where a type 1 error cannot be arbitrarily reduced without increasing the probability for a type 2 error. Another example, quite familiar to actuaries, is graduation, where a compromise between smoothness and fit has to be reached. This motivates the concept of Pareto-optimal decisions, which is discussed in section 2. There is a simple method, maximizing a weighted average of the scores, to obtain certain Pareto-optimal decisions. In section 3 a condition is given, which is satisfied in most applications, that guarantees that all the Pareto-optimal decisions can be found by this method. This is applied in section 4, where the problem of risk exchange between  $n$  insurance companies is considered. The original model of Borch is generalized: it is assumed that some of the companies are not willing to contribute more than a certain fixed amount towards the aggregate loss of the other companies. The theorem in section 4 gives a characterization of all the Pareto-optimal risk exchanges. Because of the restrictions, these risk exchanges do not just depend on the combined surplus (which would amount to pooling) in general, and can be found by an algorithm. One benefit of this generalization of Borch's Theorem is that two seemingly unrelated results (optimality of a stop loss contract, and optimality of certain dividend formulas in group insurance) follow from it as special cases.

## 2. EVALUATION OF DECISIONS UNDER CONFLICTING VIEW POINTS

Often one is faced with the situation where a decision has to be made in the presence of several criteria. Mathematically, the problem can be formulated as follows.

Let  $D$  be the set of all possible decisions. We are given  $n$  real-valued functions  $s_1(d), \dots, s_n(d)$ ,  $d \in D$ . If  $d_1, d_2 \in D$  and  $s_i(d_1) \geq s_i(d_2)$ , this means that decision  $d_1$  is *better* than (or at least as good as) decision  $d_2$  with respect to criterion  $i$ . Let

$$s(d) = (s_1(d), \dots, s_n(d)), \quad d \in D \quad (1)$$

and

$$S = \{x/x = s(d) \text{ for some } d \in D\} \quad (2)$$

denote the range of the "score function"  $s(\cdot): D \rightarrow R^n$ . A decision  $d_1$  is said to be *strictly better* than a decision  $d_2$ , if  $s_i(d_1) \geq s_i(d_2)$  for  $i = 1, \dots, n$ , and if at least one of these inequalities is strict. A decision  $d$  is called *Pareto-optimal*, if there is not a decision that

is strictly better than  $d$ . If  $R$  is any subset of  $R^n$ , a point  $x \in R$  is called a Pareto-optimal point of  $R$  if the intersection of  $R$  with  $Q_x = \{y/y_i \geq x_i, i = 1, \dots, n\}$  consists only of the point  $x$ . Thus a decision  $d$  is Pareto-optimal, if and only if  $s(d)$  is a Pareto-optimal point of  $S$ .

Under fairly general conditions (for example if  $S$  is finite, or if  $S$  is a closed region that is bounded by a plane whose normal vector points to the positive  $2^n$ -tant) one should obviously chose a Pareto-optimal decision. However, we shall not discuss the question, *which* of the Pareto-optimal decisions should be chosen.

### Example 1

In a class of  $k$  students  $n$  quizzes were given during the term. Let  $s_i(d)$  denote the score of student  $d$  in quiz  $i$  ( $i = 1, \dots, n, d = 1, \dots, k$ ). Who is the top student of the class? Thus  $D = \{1, \dots, k\}$ , and clearly the Pareto-optimal students (and only these) are candidates for this honor.

### Example 2

Consider the following statistical decision problem: Population  $i$  has a pdf  $f(x; i)$ ,  $i = 1, \dots, n$ . Given an observation, say  $X$ , the statistician tries to name the underlying population. Thus  $D$  consists of all "tests" (see [5] for example). It is convenient to allow randomized tests. Then a test  $\delta$  is defined by  $n$  non-negative functions  $p_1(x), \dots, p_n(x)$  with  $p_1(x) + \dots + p_n(x) = 1$  for all  $x$ . This means that the statistician, having observed  $X$ , will name population  $i$  with probability  $p_i(X)$ . Let

$$s_i(\delta) = \int p_i(x)f(x; i) dx \quad (3)$$

be the probability for a correct guess if the observation originates from population  $i$ ,  $i = 1, \dots, n$ . Clearly, the statistician wants to select a test that is Pareto-optimal.

### Example 3

Consider the Whittaker-Henderson Problem. Given are  $m$  ungraduated values, say  $v_1, \dots, v_m$ . A decision is the choice of  $m$  graduated values,  $u = (u_1, \dots, u_m)$ . Thus  $D = R^m$  in this example.

Let

$$F(u) = \sum_{i=1}^m w_i(u_i - v_i)^2 \quad (4)$$

be a measure for "fit", where  $w_1, \dots, w_m$  are certain weights, and let

$$S(u) = \sum_{i=1}^{m-z} (\Delta^z u_i)^2 \quad (5)$$

be our measure for "smoothness", where  $z < m$  is some integer, see [6]. Here  $n = 2$ ,  $s_1(u) = -F(u)$ ,  $s_2(u) = -S(u)$ , and we want to find graduated values that are Pareto-optimal in this sense.

The most important example (at least as far as this paper is concerned) will be discussed in section 4.

### 3. HOW TO FIND PARETO-OPTIMAL DECISIONS

Certain Pareto-optimal decisions can be found by the following method: choose  $n$  positive numbers  $k_1, \dots, k_n$  and try to maximize the linear combination

$$\sum_{i=1}^n k_i s_i(d), \quad d \in D. \quad (6)$$

For, if a decision  $\tilde{d}$  has the property that there are positive constants  $k_1, \dots, k_n$  such that

$$\sum_{i=1}^n k_i s_i(d) \leq \sum_{i=1}^n k_i s_i(\tilde{d}) \quad (7)$$

for all  $d \in D$ , it is obviously Pareto-optimal.

In *Example 1* above this method amounts to assigning certain weights to the  $n$  quizzes, and (based on this) to determine the student(s) with the highest (weighted) average score.

In *Example 2* let

$$M(x) = \max\{k_i f(x; i) / i \mid i = 1, \dots, n\}, \quad (8)$$

and let  $\tilde{\delta}$  be a test, described by  $\tilde{p}_1(x), \dots, \tilde{p}_n(x)$ , such that

$$\tilde{p}_i(x) = 0 \text{ whenever } k_i f(x; i) < M(x), \quad (9)$$

$i = 1, \dots, n$ . Thus  $\tilde{\delta}$  consists of naming the population (or one of the populations), for which the maximum is attained in formula (8).

Then if  $\delta$  is another test, given by  $p_1(x), \dots, p_n(x)$ ,

$$\begin{aligned} \sum_{i=1}^n k_i s_i(\delta) &= \sum_{i=1}^n \int k_i p_i(x) f(x; i) dx \\ &\leq \sum_{i=1}^n \int M(x) p_i(x) dx \\ &= \int M(x) dx = \sum_{i=1}^n k_i s_i(\tilde{\delta}). \end{aligned} \quad (10)$$

Hence a test  $\tilde{\delta}$  of this form is Pareto-optimal. Note that the inequality is strict unless  $\delta$  satisfies condition (9) too.

In *Example 3* the vector  $\tilde{u}$  which minimizes  $k_1F(u) + k_2S(u)$  is found as the solution of a certain matrix equation, see [6].

The question arises whether all the Pareto-optimal decisions can be obtained by this method. In general, the answer is no. Consider *Example 1* with a class of just three students. Suppose the scores in 2 quizzes were (6, 1) for student *A*, (3, 3) for student *B*, and (1, 6) for student *C*. Obviously, all 3 students are Pareto-optimal. But only students *A* and *C* can be obtained by the above method.

However, if  $S$  is a closed convex region, all the Pareto-optimal points and decisions can be obtained by this method: if  $\tilde{d}$  is a Pareto-optimal decision, inequality (7) holds for all  $d \in D$ , where  $(k_1, \dots, k_n)$  is a vector that is perpendicular to the (or a) plane that is tangent to  $S$  at  $x = s(d)$ . A convenient way to verify convexity of  $S$  is to show that for any two points  $x_0, x_1 \in S$ , the line segment  $\{x/x = rx_1 + (1-r)x_0, 0 < r < 1\}$  is contained in  $S$ . The validity of this condition can be easily seen in *Example 2*: if  $\delta_0, \delta_1$  are any two tests, define a test  $\delta_r$  ( $0 < r < 1$ ), which consists of using  $\delta_1$  with probability  $r$  and  $\delta_0$  with probability  $1 - r$ . Then, by the law of total probability,

$$s_i(\delta_r) = rs_i(\delta_1) + (1-r)s_i(\delta_0) \quad (11)$$

( $i = 1, \dots, n$ ). Hence all the Pareto-optimal tests are of the form (9), which is essentially the content of the *lemma of Neyman-Pearson*, see [5] for example.

Often it is possible to show the validity of the following condition (which may hold even if  $S$  is not convex).

*Condition C.* For any two decisions  $d_0, d_1 \in D$  there is a family of decisions  $d_r \in D$ ,  $0 < r < 1$ , such that

$$s_i(d_r) \geq rs_i(d_1) + (1-r)s_i(d_0) \quad (12)$$

for  $i = 1, \dots, n$ .

If  $S$  is closed and if Condition *C* is satisfied, all the Pareto-optimal points and decisions can be obtained by the method described at the beginning of this section: Condition *C* implies that the set of Pareto-optimal points on  $S$  coincides with the set of Pareto-optimal points on the convex hull of  $S$ . In *Example 3* the validity of Condition *C* can be verified as follows. If  $u^{(0)}, u^{(1)}$  are two vectors of graduated values

$$u^{(j)} = (u_1^{(j)}, \dots, u_m^{(j)}), j = 0, 1, \quad (13)$$

set  $u^{(r)} = ru^{(1)} + (1-r)u^{(0)}$ . Then one uses the inequality

$$(ra + (1-r)b)^2 \leq ra^2 + (1-r)b^2, \quad 0 < r < 1, \quad (14)$$

which is valid for any two numbers  $a$  and  $b$ , to show that

$$F(u^{(r)}) \leq rF(u^{(1)}) + (1-r)F(u^{(0)}) \quad (15)$$

and

$$S(u^{(r)}) \leq rS(u^{(1)}) + (1-r)S(u^{(0)}). \quad (16)$$

Therefore, all the Pareto-optimal graduated sets are obtained by the usual Whittaker-Henderson procedure, i.e., minimizing  $k_1F(u) + k_2S(u)$ .

#### 4. THE PROBLEM OF RISK EXCHANGE

Consider  $n$  insurance companies whose surplus at the end of the year will be  $X_1, \dots, X_n$ , respectively. These are  $n$  random variables with known joint distribution. The decision to be made is the selection of a risk exchange. A risk exchange is best characterized by its effect on the distribution of the surplus among the  $n$  companies. In this sense a risk exchange is a random vector

$$Y = (Y_1, \dots, Y_n), \quad (17)$$

where  $Y_i$  should be interpreted as the modified surplus of company  $i$  at the end of the year. Since the combined surplus before and after the exchange is the same, we must have

$$Y_1 + \dots + Y_n = X_1 + \dots + X_n. \quad (18)$$

We want to allow for the possibility that some of the companies are not willing to pay more than a certain amount towards the losses of the other companies. For this purpose assume  $n$  constants  $c_1, \dots, c_n$  with  $0 \leq c_i \leq \infty$ . Then only risk exchanges are admissible for which

$$Y_i \geq X_i - c_i, \quad i = 1, \dots, n. \quad (19)$$

We shall exclude the case where  $c_1 = \dots = c_n = 0$ , because in that case only the trivial "exchange" (no exchange) is possible. To summarize, a risk exchange is a random vector of the form (17) that satisfies conditions (18) and (19) with probability one.

To evaluate the different risk exchanges, assume  $n$  utility functions  $u_1(x), \dots, u_n(x)$ ,  $-\infty < x < \infty$ . Suppose that these functions are twice differentiable, with

$$u'_i(x) > 0, \quad u''_i(x) \leq 0. \quad (20)$$

For simplicity, we shall also assume that at most one of these utility functions is linear and that all of the others have the prop-

erty that their derivative decreases from  $\infty$  to 0 as the argument increases from  $-\infty$  to  $\infty$ . Then a problem of the following type has a unique solution: given a number  $\lambda$  and positive numbers  $k_1, \dots, k_n$ , find numbers  $z_1, \dots, z_n$  such that

$$k_i u_i'(z_i) \text{ is independent of } i \quad (21)$$

and

$$z_1 + \dots + z_n = \lambda. \quad (22)$$

This solution  $z = (z_1, \dots, z_n)$  has a geometric interpretation: it corresponds to the point on the surface

$$F_\lambda = \{x = (x_1, \dots, x_n) / x_i = u_i(t_i), t_1 + \dots + t_n = \lambda\} \quad (23)$$

where the tangential plane is perpendicular to the vector  $(k_1, \dots, k_n)$ . In the case of exponential utility functions,

$$u_i(x) = \alpha_i(1 - \exp(-x/\alpha_i)), \quad (24)$$

where  $x_1 > 0, \dots, x_n > 0$ , this problem can be solved explicitly. One finds that

$$z_i = \beta_i \lambda + \alpha_i (\log k_i - \sum_{j=1}^n \beta_j \log k_j), \quad (25)$$

where  $\beta_i = \alpha_i / (x_1 + \dots + x_n)$ .

It is assumed that company  $i$  is only interested in the expected utility of its own surplus,

$$s_i(Y) = E[u_i(Y_i)], \quad (26)$$

$i = 1, \dots, n$ . In this sense we are faced with the problem of finding Pareto-optimal risk exchanges. Let us verify the validity of Condition C in this case. If  $Y^{(0)}, Y^{(1)}$  are any two risk exchanges,

$$Y^{(j)} = (Y_1^{(j)}, \dots, Y_n^{(j)}), j = 0, 1, \quad (27)$$

define

$$Y^{(r)} = (Y_1^{(r)}, \dots, Y_n^{(r)}), 0 < r < 1, \quad (28)$$

by setting

$$Y_i^{(r)} = rY_i^{(1)} + (1-r)Y_i^{(0)}. \quad (29)$$

Since  $Y^{(0)}$  and  $Y^{(1)}$  satisfy (18) and (19), it follows that  $Y^{(r)}$  satisfies these conditions. Thus  $Y^{(r)}$  is a risk exchange. Since the function  $u_i$  is concave from below,

$$u_i(Y_i^{(r)}) \geq r u_i(Y_i^{(1)}) + (1-r) u_i(Y_i^{(0)}). \quad (30)$$

Taking expected values, we get

$$s_t(Y^{(r)}) \geq r s_t(Y^{(1)}) + (1-r) s_t(Y_t^{(0)}), \quad (31)$$

which shows that Condition C holds. Obviously,  $S$  is closed, so to find the Pareto-optimal risk exchanges it is enough to choose positive constants  $k_1, \dots, k_n$  and to try to maximize

$$\sum_{i=1}^n k_i s_t(Y) = \sum_{i=1}^n k_i E[u_t(Y_i)]. \quad (32)$$

In this paragraph we shall construct a risk exchange  $Y$  and then verify that it maximizes (32). Let

$$I^{(0)} = \{1, \dots, n\}, J^{(0)} = \emptyset. \quad (33)$$

We define random vectors

$$Z^{(m)} = (Z_1^{(m)}, \dots, Z_n^{(m)}) \quad (34)$$

and index sets  $I^{(m)}$  and  $J^{(m)}$  as follows. For  $m = 1, 2, \dots$  set

$$Z_i^{(m)} = X_i - c_i \text{ if } i \in J^{(m-1)} \quad (35)$$

and choose  $Z_i^{(m)}, i \in I^{(m-1)}$ , such that

$$k_i u'_i(Z_i^{(m)}), i \in I^{(m-1)}, \text{ is independent of } i \quad (36)$$

and

$$\sum_{i \in I^{(m-1)}} Z_i^{(m)} = \sum_{i \in I^{(m-1)}} X_i + \sum_{j \in J^{(m-1)}} c_j. \quad (37)$$

Then

$$I^{(m)} = \{i | Z_i^{(m)} > X_i - c_i\} \quad (38)$$

and

$$J^{(m)} = \{i | Z_i^{(m)} \leq X_i - c_i\}. \quad (39)$$

From this recursive definition it follows immediately that

- (i)  $I^{(m)} \subseteq I^{(m-1)}, J^{(m)} \supseteq J^{(m-1)}$
- (ii)  $I^{(m)}$  is not empty.

Furthermore, if  $M^{(m)}$  denotes the common value of the expressions in (36), one can show that

- (iii)  $M^{(m+1)} \geq M^{(m)}$
- (iv)  $k_i u'_i(Z_i^{(m)}) \leq M^{(m)}$
- (v)  $k_i u'_i(Z_i^{(m)}) < M^{(m)}$  implies  $Z_i^{(m)} = X_i - c_i$ .

Now let  $\tilde{Y} = (\tilde{Y}_1, \dots, \tilde{Y}_n)$  be the limit of  $Z^{(m)}, m \rightarrow \infty$ . (Note

that this limit is obtained after finitely many steps; as a matter of fact,  $\tilde{Y} = Z^{(n)}$ .) Observe that  $\tilde{Y}$  is a risk exchange and has the following property:

*Property B*

Let  $M = \max\{k_i u'_i(\tilde{Y}_i) / i = 1, \dots, n\}$ . Then  $k_i u'_i(\tilde{Y}_i) < M$  implies that  $\tilde{Y}_i = X_i - c_i$ .

We shall now compare  $\tilde{Y}$  with an arbitrary risk exchange  $Y = (Y_1, \dots, Y_n)$  as follows: since the function  $u_i(\cdot)$  is concave from below, and since  $k_i u'_i(Y_i) < M$  implies that  $Y_i \geq \tilde{Y}_i$ ,

$$\begin{aligned} k_i u_i(Y_i) &\leq k_i u_i(\tilde{Y}_i) + k_i u'_i(\tilde{Y}_i) \cdot (Y_i - \tilde{Y}_i) \\ &\leq k_i u_i(\tilde{Y}_i) + M \cdot (Y_i - \tilde{Y}_i). \end{aligned} \quad (40)$$

Thus

$$\sum_{i=1}^n k_i u_i(Y_i) \leq \sum_{i=1}^n k_i u_i(\tilde{Y}_i) \quad (41)$$

and

$$\sum_{i=1}^n k_i E[u_i(Y_i)] \leq \sum_{i=1}^n k_i E[u_i(\tilde{Y}_i)]. \quad (42)$$

Furthermore, the last inequality is strict unless  $Y = \tilde{Y}$  (almost surely). Our findings can be summarized as follows.

*Theorem*

a) Given  $k_i > 0, \dots, k_n > 0$ , there is exactly one risk exchange that satisfies Property B. b) A risk exchange is Pareto-optimal if and only if it is of this form.

*Special cases*

1) If  $c_1 = \dots = c_n = \infty$ , this result reduces to the classical *Theorem of Borch*, see [2], [3], or [4].

2) Consider the case, where  $u_1(x) = x$ ,  $c_1 = \infty$ ,  $u_2(x) = u(x)$  (strictly concave from below), and  $c_2 = P > 0$ . We find that the Pareto-optimal risk exchanges are of the form

$$\tilde{Y}_1 = \begin{cases} X_1 + P & \text{if } X_2 \geq \alpha \\ X_1 + P - (\alpha - X_2) & \text{if } X_2 < \alpha \end{cases} \quad (43)$$

$$\tilde{Y}_2 = \begin{cases} X_2 - P & \text{if } X_2 \geq \alpha \\ \alpha - P & \text{if } X_2 < \alpha \end{cases} \quad (44)$$



where the parameter  $\alpha$ , satisfying the equation  $k_1 = k_2 u'(\alpha - P)$ , plays the role of a deductible. This result (*optimality of a stoploss contract*) is due to Arrow, see [1].

3) Consider the case, where  $u_1(x) = x$ ,  $c_1 = 0$ ,  $u_2(x) = u(x)$  (strictly concave from below), and  $c_2 = \infty$ . Thus  $Y_1 = X_1 + D$ ,  $Y_2 = X_2 - D$ , where  $D \geq 0$  is a *dividend* payable from company 2 to company 1. We find that Pareto-optimal dividends are of the form

$$\tilde{D} = \begin{cases} X_2 - \alpha & \text{if } X_2 > \alpha \\ 0 & \text{if } X_2 \leq \alpha \end{cases} \quad (45)$$

This result has been found in [7] in connection with dividend formulas in group insurance.

#### REFERENCES

- [1] ARROW, K., (1963), "Uncertainty and the Welfare Economics of Medical Care", *American Economic Review*, LIII, 941-73.
- [2] BORCH, K., (1962), "Equilibrium in a reinsurance market", *Econometrica* 30, 424-44
- [3] BÜHLMANN, H., 1970, *Mathematical Methods in Risk Theory*. New York: Springer
- [4] DUMOUCHEL, W. H., (1968), "The Pareto-optimality of a n-company reinsurance treaty", *Skandinavisk Aktuarietidskrift* 51, 165-70
- [5] DEGROOT, M. H., 1975 *Probability and Statistics*, Reading: Addison-Wesley.
- [6] GREVILLE, T. N. E., (1973), *Part Five Study Notes on Graduation*, Chicago: Society of Actuaries.
- [7] JONES, D. A. and GERBER, H. U., (1974), "Dividend Formulas in Group Insurance", *Transactions of the Society of Actuaries* 26, 77-93.