COWLES COMMISSION DISCUSSION PAPER: ECONOMICS NO. 2074

"The Best Can Be An Enemy of the Good." [1]

L. Törnqvist

In the theory of rational decisions we usually try to find optimal decisions among a set of possible decisions. A decision maker who is very anxious to make only optimal decisions can however be beaten in a contest by a decision maker who makes many more decisions per unit of time and deliberately takes the risk of making some non-optimal decisions thereby saving so much decision making time that he can compensate the loss due to the nonoptimality of single decisions by means of a large number of approximately optimal decisions. If the decision maker gets valuable information because he has departed from the way of optimal decisions, this fact can in the long run more than compensate for the direct losses that can be assigned to a "bad" decision.

The strategy "always make the best decision possible" can thus be nonoptimal in the long run. The truly optimal strategy contains a large number of approximately optimal decisions. The distribution of the decisions according to the difference between the value of the optimal decision and the value of the decision actually made will thus in general

---

1/ Research undertaken by the Cowles Commission for Research in Economics under contract Nonr-358(01), N F 047-006 with the Office of Naval Research.

have a positive variance. Occasionally some bad decisions have to be tolerated to make it possible for the decision maker to increase the speed of decision making. By studying the consequences of bad decisions which have been made, it is often possible to get valuable information about the value of future prospects by different actions. This information can later on make it easier to find good decisions than if no "bad" decision had been made. If we learn from our mistakes, the mistakes can be worth their price.

These common experiences can be formalized in the following way: Let $F_1(z) = P(U_a - U_i \leq z)$ denote the probability to make a decision which is at most z utility units worse than the optimal decision possible in the situation considered; and let $F_2(x, n; m)$ denote the probability that during the time period considered situations will occur which are such that the corresponding optimal decision will be $U_\alpha \geq x > 0$ utility units better than the wait and see action $i = 0$. Depending on the time spent on each decision the decision maker is not able to make more than a finite number of decisions $i \neq 0$. If he tries to behave in such a way that

$$\int z \, dF_1(z) = m$$

the expected number of decisions he is able to make might be $N(m)$. If he makes $N(m)$ decisions he can get at most $\bar{U}(m) = \int_{x \in X_m} x dF_2(x, N(m); m)$ units per decision $i \neq 0$ if he makes only optimal decisions in each case. When he makes approximately optimal decisions his mean gain is $\bar{U}(m) - m$. In all, his decision making activity will during the period considered give him a profit

$$V_m = N(m) \; (\bar{U}(m) - m).$$

This profit is maximized if m is such that

$$\frac{\partial \log N(m)}{\partial m} = - \frac{\partial \log(\bar{\bar{U}}(m) - m)}{\partial m}$$

that is, if the relative increase in the expected number of decisions ($i \neq 0$) is equal to the relative decrease in the mean profit $\bar{\bar{U}}(m)-m)$ per decision he makes, if he with the purpose to be able to make more decisions deliberately decreases the quality measured by the mean loss m due to nonoptimality of each single decision. By using the strategy to make the decision $i = 0$ in all cases when U is smaller than some number $x_o > 0$ and some other case which needs his attention is waiting, he can concentrate his decision making activity to those situations which give a relatively high mean value to $\bar{\bar{U}}(m)$. In cases when the probability for losses $z > m$ due to nonoptimality is comparatively large it pays to use more time to these decisions than to cases where this probability is small.

The possibility, that nonoptimal decisions can be a source of more valuable information for future decision making than optimal decisions, can be taken into account formally by deducting the value of such information from the quantity m.

In many cases the need to make many decisions for unit of time is a result of nonoptimal decisions made earlier. The decision maker has namely often to make many new decisions to minimize the bad consequences of some decision already made. If the loss due to nonoptimality is large this can result in a situation where a large part of the decision making activity has to be used for such decisions which had been unecessary to make if more attention had been paid to the first decision in the chain of mutually connected decisions.

It is thus a good policy not to be too anxious to make only optimal decisions, but decisions which are not even approximately optimal has to be avoided, whenever possible.

If we will avoid the paradox that the truly optimal strategy does not necessarily consists of a sequence of optimal single decisions, we have to develop the theory of the rational single decisions by taking in account the fact that the decision making activity istelf is not costless. Such a theory could perhaps be constructed by means of the following ideas.

The Probability of Rational Decisions.

Because the problem of finding optimal solutions to decision-making problems usually is very difficult and requires efforts, that is costs, which have to be included as a part of the anticipated consequences influencing the decisionmaking activity, the decisionmaker will in practice make his choices to some extent at random between the alternatives considered. We shall therefore study the probabilities of different choices under the assumption that they are functions of the values assigned to the set of anticipated consequences. We will find that there exists an optimal degree of randomness in the final decisions depending on the costs of finding good solutions. This degree of randomness left is nearly connected with the entropy concept. We will also get an explanation of why we by practical experiments will sometimes get results that seem to contradict the assumption that the values U_i attached to the anticipated consequences of choices considered constitute a well ordered set.

<u>We will use the following notations:</u>

(1) $a_i \in A$ = actions possible for the decision maker

(2) $H_i \in H$ = prospects of consequences by choice of a_i

(3) $H_{iz} \in H_i$ = possible consequences of the choice a_i

(4) $V_i = V(H_i) = V(\underset{z}{E}\, U\,(H_{iz}))$ = Indicator of goodness for a_i given the sets H and A, $U_{iz} = U(H_{iz})$ utility of H_{iz}, U_i = expected value of U_{iz}.

(5) C = an indicator of efforts made with the goal to find an optimal action a_α.

(6) $\hat{C}$ = Optimal efforts for the purpose to find α given H, A, usually increasing with the complexity of (H, A).

(7) $\gamma \approx C \mid \hat{C}$. C_1 a parameter influencing the probabilities to choose a_i as the final decision.

(8) p_i = the probability of the choice a_i given A, H, γ.

(9) $(V_i \geq V_j) \Longleftrightarrow (p_i \geq p_{j\gamma})$ <u>weak postulate</u> of rationality.

We will strengthen this weak postulate a little by assuming that the odds p_i/p_j are functions of the difference $V_i - V_j$. This assumption is also in fact very weak if we observe that the values $V_i = V(U_i) = V(U(H_i))$, where $V(U)$ is a monotonically increasing function of the primary values U_i assigned to H_i. If

$$\frac{p_i}{p_j} = \frac{p_i}{p_k} \cdot \frac{p_x}{p_j} = P(V_i - V_j) = P(V_i - V_k) \cdot P(V_k - V_j)$$

the linearity condition:

$$\log P(V_i - V_j) = \log P(V_i - V_k) + \log P(V_k - V_j) = \gamma.(V_i - V_j)$$

(compare $L(x + y) = L(x) + L(y)$; $L(x) = \log P(x))$ only if $L(x) = \gamma x$.)

must be fulfilled. We get

(10) $\frac{p_i}{p_j} = e^{\gamma(V_i - V_j)} = e^{\gamma(V(U_i) - V(U_j))}$ = strong postulate about the relation between p_i/p_j and $V_i - V_j$. From the week postulate follows $\gamma \geq 0$.

(11) $V_i = V_j$ if $p_{i\gamma} = p_{j\gamma}$,

further:

(12) $p_{i\gamma} = e^{\gamma(V_i - V_\alpha)} \Big/ \sum_j e^{\gamma(V_j - V_\alpha)}$; $\gamma > 0$

and

(13) $p_{\alpha\gamma} \longrightarrow 1,\ p_{j\gamma} \longrightarrow 0,\ j \neq \alpha$ if $\gamma \longrightarrow \infty$,

$V_\alpha = \max_{a_i \in A} V_i > V_j$ for every $j \neq \alpha$.

The formula

(14) $\gamma(V_\alpha - V_i) = \log\ p_{\alpha\gamma} - \log\ p_{i\gamma}$,

defines a metric for V_i.

(15) $R\gamma = \Sigma\, p_{i\gamma}(V_\alpha - V_i) \longrightarrow 0, \qquad \gamma \longrightarrow \infty$

R_γ = expected regret when the efforts $C = \gamma \hat{C}\ C_1$ are made to find α. If C is measured by the decrease in V_α, when the value of the efforts to find α are deducted from the value of the prospect H_α, the optimal efforts $\hat{C}$ must be such that "marginal revenue = marginal costs," that is

(16) $-\frac{\partial R_\gamma}{\partial \gamma} = \frac{\partial C}{\partial \gamma} = \hat{C}\ C_1$ for $\hat{C} = \gamma C_1\ \hat{C},\ \gamma = 1/\ C_1 = \hat{\gamma}$

It follows that for $\gamma = \hat{\gamma}$ $\qquad \gamma = \frac{C\,\hat{\gamma}}{\hat{C}}$.

$$\hat{C}\, C_1 = \sum_i p_{i\hat{\gamma}}(V_i - V_\alpha)^2 - R^2_{\hat{\gamma}} = \hat{\sigma}^2$$

and consequently

$$\gamma = \frac{C}{\hat{\sigma}^2}$$

$\hat{\sigma}^2$ is a measure of variance for optimal decisions. The parameter $\hat{C}/\hat{\sigma}^2$ is a characteristic of the decision makers "skill" to make with little effort C good decisions.

The "odds" for the decision a_α compared with the odds for the decision a_i can for optimal effort $\hat{C}$ be written

$$(17)\quad \frac{p_{a\hat{\gamma}}}{p_{i\hat{\gamma}}} = e^{\hat{C}(V_\alpha - V_i)/\hat{\sigma}^2} = e^{\hat{\gamma}(V_\alpha - V_i)}$$

The "odds" for the decision a_α increases rapidly with the skill $\hat{\gamma} = \hat{C}/\hat{\sigma}^2$ and the difference between the goodness of a_α and a_i.

The probability that the alternative a_i would be chosen n_i times out of n decision when the decisionmaker chooses only between the alternatives i and j can be calculated according to the binomial-formula:

$$p_j(n_i;\, n) = \binom{n}{n_i}\, p_i^{n_i}\, p_j^{n_j} \Big/ (p_i + p_j)^n$$

The probability that $n_i \geq \frac{n}{2}$ is then

$$P(n_j \leq n_i) = \sum_{n_j \leq \frac{n}{2}} \binom{n}{n_j}\, p_i^{n_i}\, p_j^{n_j} \Big/ (p_i + p_j)^n$$

$$\approx \Phi\left(\frac{\frac{n}{2}(p_i + p_j) - n\, p_j}{\sqrt{n\, p_i\, p_j}}\right) = \Phi\left(\sqrt{\frac{n}{2}}\left(\sqrt{\frac{p_i}{p_j}} - \sqrt{\frac{p_j}{p_i}}\right)\right)$$

If we observe that

$$\frac{1}{2}\left(\sqrt{\frac{p_i}{p_j}} - \sqrt{\frac{p_j}{p_i}}\right) = \frac{1}{2}\left(e^{\frac{\gamma(V_i - V_j)}{2}} - e^{\frac{-\gamma(V_i - V_j)}{2}}\right) = \sinh \frac{\gamma}{2}(V_i - V_j)$$

we get $P(n_j \leq n_i) \approx \Phi(\sqrt{n} \sinh \frac{\gamma}{2}(V_i - V_j)) = \Phi_{ij}(n)$

$\Phi(x)$ denote the normal distribution $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{\frac{-x^2}{2}} dx$

If $\gamma(V_i - V_j)$ is not very small this probability increases rapidly towards 1 when n increases.

If we make three independent trials to determine if i is preferred to j and j to k and i to k when $V_i > V_j > V_k$ we find that the probability of the transitive relation $n_i \geq \frac{n}{2}$, $n_j > \frac{n}{2}$, $n_k > \frac{n}{2}$ is approximately equal to:

$$P(n_i \geq \frac{n}{2}, n_j \geq \frac{n}{2}, n_k \geq \frac{n}{2}) = \Phi_{ij}(n) \cdot \Phi_{jk}(n) \cdot \Phi_{ik}(n).$$

This probability increases also towards 1 when $n \longrightarrow \infty$ if $V_i > V_j > V_k$

The rapidity of the convergence depends almost wholly on the smallest difference between these values.

If the three alternatives are simultanously available we get the probabilities:

$$P(n_i, n_j, n_k; n) = \frac{n!}{n_i!\, n_j!\, n_k!} p_i^{n_i} p_j^{n_j} p_k^{n_k}$$

for the triplet (n_i, n_j, n_k) ; $n = n_i + n_j + n_k$. The most probable triplet $\hat{n}_i, \hat{n}_j, \hat{n}_k$ is such that if λ, ρ denotes a pair from (i, j, k)

$$\log(\hat{n}_\lambda + \theta_\lambda) - \log(\hat{n}_\rho + \theta_\rho) = \log p_\lambda - \log p_\rho = \gamma(V_\lambda - V_\rho)$$

and

$$\frac{p_i V_i + p_j V_j + p_k V_k}{p_i + p_j + p_k} \qquad \frac{n_i V_i + n_j V_j + n_k V_k}{n}$$

where $|\theta_\lambda| = 1$, $|\theta_\rho| = 1$.

In Shannon's theory of information he comes to the conclusion that a concept corresponding to the entropy of a system is of fundamental importance for such a theory. This entropy is a quantity of the form $W_\gamma = -\sum_i p_{i\gamma} \log p_{i\gamma}$. W_γ is a measure randomness of the decisionmaking for given γ.

If we in the formula for the expected regret $R_\gamma = \Sigma\, p_{i\gamma}(V_{\alpha\gamma} - V_{i\gamma})$ introduce $V_{\alpha\gamma} - V_{i\gamma} = \frac{1}{\gamma}\,(\log p_{\alpha\gamma} - \log p_{i\gamma})$ we get

$$R_\gamma = (\log p_{\alpha\gamma} + W_\gamma)/\gamma$$

The expected regret R_γ is thus very nearly connected with the entropy concept.

The assumption that the indicator of efforts or costs C_γ is proportional to γ is however somewhat arbitrary. The principle to minimize $R_\gamma + C_\gamma$ is perhaps also unsatisfactory. A more general approach is to assume that the values assigned to the choice a_i is a function $U_{i\gamma}$ of γ and $V_i = V_{i\gamma} = V(U_{i\gamma})$. Instead of minimizing $R_\gamma + C_\gamma$ we can substitute the principle of maximizing.

$$U_\gamma = \sum_i U_{i\gamma} p_{i\gamma} \quad ; \quad p_{i\gamma} = e^{\gamma V_{i\gamma}} \Big/ \sum_i e^{\gamma V_{i\gamma}}$$

The condition for optimal choice of γ can then be written in the form:

$$C'_\gamma = -\sum_i \frac{\partial U_{i\gamma}}{\partial \gamma} \cdot p_{i\gamma} = \sum_i U_{i\gamma} V_{i\gamma} p_{i\gamma} - \sum_i U_{i\gamma} p_{i\gamma} \sum_i V_{i\gamma} p_{i\gamma}$$

$$+ \gamma\Big(\sum_i U_{i\gamma} \frac{\partial V_{i\gamma}}{\partial \gamma}\, p_{i\gamma} - \sum_i U_{i\gamma} p_{i\gamma} \cdot \Sigma \frac{\partial V_{i\gamma}}{\partial \gamma}\, p_{i\gamma}\Big)$$

or

$$C'_{\gamma} = r_{uv;\gamma}\sigma_{v\gamma}\sigma_{u\gamma} + \gamma r_{uv';\gamma}\sigma_{v'u;\gamma}\sigma_{u\gamma}$$

where C'_{γ} corresponds to the marginal costs of increasing γ, $r_{uv;\gamma}$ denotes the coefficient of correlation between $U_{i\gamma}$ and $V_{i\gamma}$ and $r_{uv';\gamma}$ the coefficient of correlation between $U_{i\gamma}$ and $\frac{\partial V_{i\gamma}}{\partial \gamma}$, and $\sigma_{u\gamma}$, $\sigma_{v\gamma}$, $\sigma_{v'\gamma}$ standard deviation for $U_{i\gamma}$, $V_{i\gamma}$ and $\frac{\partial V_{i\gamma}}{\partial \gamma}$ respectively.

If the scaling of $U_{i\gamma}$ and $V_{i\gamma}$ is chosen so that $\sigma_{u\hat{\gamma}} = \sigma_{v\hat{\gamma}}$ we get the result:

$$C'_{\hat{\gamma}} = r_{uv;\hat{\gamma}}\sigma^2_{u\hat{\gamma}} + \hat{\gamma} r_{uv';\hat{\gamma}}\sigma_{v'\hat{\gamma}}\sigma_{u\hat{\gamma}}$$

In the case when $\gamma r_{uv';\hat{\gamma}}\sigma_{v'\hat{\gamma}}$ is small $C'_{\hat{\gamma}}$ is approximately equal to $r_{uv;\hat{\gamma}}\sigma^2_{u\hat{\gamma}}$. If further, $U_{i\gamma}$ and $V_{i\gamma}$ are linearly dependent $r_{uv;\gamma} = 1$ and

$$C'_{\hat{\gamma}} = \sigma^2_{u\hat{\gamma}} = \sigma^2_{v\hat{\gamma}}.$$

The model introduced seems to me to be so flexible that it cannot be contradicted by any empirical tests. It might however be possible to get some information about the question if $r_{uv;\gamma} = 1$ and about the manner in which the parameter γ varies between different types of decisionmakers, when they have to choose between the same alternatives.

It might be possible to study how γ is influenced by the time available for decisionmaking and the complexity of the decisionmaking in question. This quantity seem thus to be of considerable interest when we try to study the psykometric problems connected with the decisionmaking ability of different persons.