THE EFFICIENT USE OF AN IMPERFECT FORECAST*

HOWARD E. THOMPSON AND WILLIAM BERANEK

University of Wisconsin

This paper illustrates how individual forecasts and forecasting techniques may be evaluated by the use of established decision theory. Given the probability distribution of the forecast error, we first find the optimal strategy for a decision process, i.e., how to make the most efficient use of a forecast. After expressing the expected profit of the optimal strategies in terms of the probability of a correct forecast, we illustrate how to find (1) the value of an imperfect forecast, and (2) the value of a forecasting technique. With this information, we can then determine when to use a forecast, the maximum amount to pay for both a forecast and a forecasting method, and the conditions under which it is worthwhile to attempt to improve the accuracy of a forecasting method.

While substantial sums are expended on improving forecasting procedures, little attention is devoted to the questions of when to forecast, how to use a forecast, and, what is equally important, when to attempt to improve a forecasting procedure. For many situations, the tools and the empirical data necessary for answering these questions are at hand.

The purpose of this paper is to illustrate, by means of a simple three-event inventory model, how one can obtain answers to these important questions. First, we shall show how the probability of a correct forecast may be used to determine the optimal use of an *imperfect* forecast. Secondly, while the concept of the value of a *perfect forecast* is well known, we will develop and apply the notion of the value of an *imperfect forecast*. Thirdly, we shall indicate the conditions under which attempts to improve forecasting accuracy are warranted and when, surprisingly enough, a more accurate forecasting procedure, even if cost free, would be rejected in favor of a less accurate method. Finally, we shall extend these concepts to a general multi-event decision process.

In the analysis to follow we make use of the techniques and concepts of existing decision theory. We are not concerned with either *inference*, *Bayesian statistics*, or with the further development of decision theory.

The Optimal Use of a Forecast

To illustrate the optimal use of an imperfect forecast we shall assume an inventory problem in which the amounts demanded consist of 1, 2, or 3 units with probabilities .3, .4, and .3 respectively. Let x and z denote order quantities and amounts demanded respectively. If it costs \$1.50 to stock a unit, if each unit sells for \$5.00, and if there is no salvage value for unsold units, then the profit matrix is given by Table 1 where it is understood, of course, that the dimension of profit is dollars.

^{*} Received June 1965 and revised April 1966.

Projit matrix			
	s = 1	s = 2	z = 3
x = 1 $x = 2$ $x = 3$	3.50 2.00 .50	3.50 7.00 5.50	3.50 7.00 10.50

TABLE 1
Profit Matrix

If there is no forecasting the optimal policy is frequently found by evaluating the expected profit of each order quantity, $\epsilon(x)$, and then selecting that order quantity with the largest expected value. In this example we have

$$\epsilon(x = 1) = $3.50$$
 $\epsilon(x = 2) = 5.50
 $\epsilon(x = 3) = 5.50

and hence either x = 2 or x = 3 is the optimal order quantity.

Perfect Forecasts

If, however, it is possible to forecast the period's demand without error, then clearly an inspection of Table 1 reveals that the optimal amount to stock varies with the demand and corresponds to the amount of the forecast. If the forecast is for 1, 2, or 3 units, the best policy would be to order 1, 2, or 3 units respectively. When used with a perfect forecast, this would lead to an expected profit of

$$.30(\$3.50) + .40(\$7.00) + .30(\$10.50) = \$7.00$$

Since the best we can do with a perfect forecast in each period is an average long-run profit of \$7.00—a result which is often labeled the expected profit under certainty—while the largest expected profit per period in the absence of a forecast is \$5.50, the maximum amount we would expend for a perfect forecast is \$1.50. The quantity has been given a variety of appellations the most common being the expected value of perfect information, the cost of the risk, and the cost of uncertainty.

Imperfect Forecasts

While the above procedure determines the optimal policy for a perfect forecast, we must now develop a method for determining the optimal strategy for an imperfect forecast, a method in which obtaining the optimal policy in the face of a perfect forecast is simply a special case. Let us turn now to the problem of finding the optimal strategy for an imperfect forecast.

Although the quantity demanded is a random variable with a known distribution, a forecaster may have insights, or additional information, which enables him to predict this event. The amount forecast for each period will not normally correspond to the realized value, but will be distributed about this value, the

	s = 1	s = 2	s = 3
F = 1 F = 2 F = 3	$1 - \gamma$	$\frac{(1-\gamma)/2}{\gamma}$ $\frac{\gamma}{(1-\gamma)/2}$	0 1 - γ γ

TABLE 2
Conditional Probability of a Forecast, F, Given Demand, z

"state of nature." It is assumed, however, that the probability distributions of the amount of the forecast, given the realized quantities demanded, are at hand.

The reader who is acquainted with the basic elements of decision theory will recognize that a "forecast" in this framework is simply a special case of what is commonly referred to as an "experiment." The probability distribution of the outcomes of the experiment, given the state of nature, corresponds to the conditional probability distribution of the forecast in our framework, while the decision theorist's a priori probabilities of the states of nature correspond to our probabilities of the quantity demanded. The conditional probability distributions of the amount of the forecast can often be developed from the firm's records, and our main task is to evaluate the forecaster's worth in the light of these conditional distributions and the demand distribution.

While Chernoff and Moses [1, pp. 104–14] have set forth procedures for determining the worth of a proposed "experiment," we will show how an evaluation of a forecast may be made in terms of its accuracy. This, in turn, will enable us to assess an imperfect forecast as well as a proposed reduction in the forecast error.

To this end let F denote the amount of the forecast. We shall assume that the forecaster is correct with probability γ , i.e., the probability that he chooses F=z is precisely γ . Then the conditional probability distributions of F for each value of z are given by Table 2. This table, which we may call a forecast matrix, contains the probabilities of the forecast errors. The assumption that F=z with probability γ is made solely for the sake of simplification and does not affect the analysis to follow. The remaining elements of the matrix are arbitrarily selected.

Since to each possible forecast there are three acts which may be undertaken (excluding x=0), there are 3^3 or 27 different policies or strategies which must be evaluated. One strategy consists of ordering 1 unit if the forecast is 1, 2 if the forecast is 2, and 3 if the forecast is 3. Following Chernoff and Moses [1, p. 5] this particular strategy can be denoted as S(1, 2, 3), where the position of the integer denotes the amount of the forecast—i.e., for example, the first position will always indicate a forecast of 1 unit—while the integer occupying the position denotes the quantity ordered corresponding to the forecast. Thus the strategy S(2, 2, 3) would mean precisely to order 2 when the forecast is 1, 2 when the forecast is 2, and 3 when the forecast is 3.

It is now possible, given γ , to compute the expected profit of each strategy and then choose that one which maximizes expected profits. However, some

strategies can be neglected because they are obviously inferior. For example, the probability of F=1 given z=3 is zero and hence, no rule should call for an order of 3 if F=1, i.e., all strategies with the number 3 in the first position are inferior strategies and need not be evaluated. Likewise since z=1 never yields a forecast of 3 units, a strategy with the number 1 in the third position would never be optimal and therefore can be ignored. Since there are 15 rules embodying these possibilities and since these are inferior strategies we are left with 27-15 or 12 strategies to be evaluated. While each of these, which we may label as selective strategies, may be evaluated to determine the optimal strategy for given γ , it is more revealing to derive the best strategies for all $0 \le \gamma \le 1$. This will provide us with a number of useful insights.

To this end let us evaluate the strategy S(1, 2, 3). If z = 1, then 1 unit will be ordered with probability γ and 2 will be ordered with probability $1 - \gamma$. Thus the average profit will be $3.50 \ \gamma + 2.00(1 - \gamma) + .50(0)$. If z = 2, then 1 unit will be ordered with probability $(1 - \gamma)/2$, 2 will be ordered with probability γ , and 3 will be ordered with probability $(1 - \gamma)/2$. Hence when z = 2 average profit becomes $3.50(1 - \gamma)/2 + 7.00 \ \gamma + 5.50(1 - \gamma)/2$. Finally, when z = 3 we have an average profit of $3.50(0) + 7.00(1 - \gamma) + 10.50(\gamma)$. The expected profit from following the rule S(1, 2, 3), denoted as $\epsilon[S(1, 2, 3)]$, is then $\epsilon[S(1, 2, 3)] = 2.50 \ \gamma + 4.50$, a linear function of the probability of a correct forecast, γ . In fact, as will be shown, under these conditions each of the remaining 11 strategies is likewise a linear function of the probability of a perfect forecast.¹

Evaluating these 11 strategies in like manner we emerge with expected profit as a function of γ for each strategy and hence

ю	a function of 7 for each strategy and h	EHCE
	$(1) \ \epsilon[S(1, 2, 3)] = 2.50 \ \gamma + 4.50$	$(7) \ \epsilon[S(2, 1, 2)] =80 \ \gamma + 4.90$
	$(2) \ \epsilon[S(1, 1, 2)] = .35 \ \gamma + 4.20$	(8) $\epsilon[S(2, 1, 3)] = .55 \gamma + 4.60$
	$(3) \ \epsilon[S(1, 1, 3)] = 1.70 \ \gamma + 3.90$	$(9) \ \epsilon[S(2, 2, 2)] = 5.50$
	$(4) \ \epsilon[S(1, 2, 2)] = 1.15 \ \gamma + 4.80$	$(10) \ \epsilon[S(2, 2, 3)] = 1.35 \ \gamma + 5.20$
	$(5) \ \epsilon[S(1, 3, 2)] =05 \ \gamma + 5.40$	$(11) \epsilon[S(2, 3, 2)] = -1.20 \gamma + 6.10$
	(6) $\epsilon[S(1, 3, 3)] = 1.30 \gamma + 5.10$	$(12) \ \epsilon[S(2, 3, 3)] = .15 \ \gamma + 5.80$

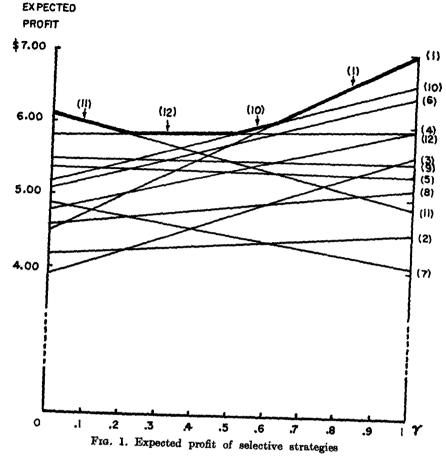
The Optimal Strategies for the Forecasting Case

The above expected profit functions for the 12 selective strategies are plotted in Figure 1. Since the 15 obviously inferior strategies yield, for any γ , an expected profit below the expected profit of at least one of the selective strategies, the optimal strategy for the forecasting case is a selective strategy which depends clearly on γ . For $\gamma=0$, the optimal strategy is S(2,3,2), while for $\gamma=1$ it is S(1,2,3), the result which emerges from perfect forecasts. For $0<\gamma<1$, it is seen that the locus of points denoting the expected profits of the optimal strategies consists of the connected heavy line segments.

Uncertainty of the Value of γ

In the event γ is not known with certainty, the sensitivity of the choice of a

A more efficient computational procedure for evaluating strategies is set forth later.



strategy to changes in γ may be assessed easily. If γ is known to lie in the interval $\gamma_0 < \gamma < \gamma_1$, for example, then determine the optimal strategies associated with γ_0 and γ_1 . If these strategies are identical, then the optimal strategy is independent of $\gamma_0 < \gamma < \gamma_1$ and our lack of knowledge concerning the exact value of γ is of no concern. Should the strategies not correspond and if no strategy other than the optimal strategies associated with γ_0 and γ_1 is optimal in the interval $\gamma_0 < \gamma < \gamma_1$, then one can choose the strategy which minimizes the expected loss for failure to employ the correct strategy. This is determined as follows.

Let $\epsilon(S_1, \gamma_0)$ and $\epsilon(S_2, \gamma_1)$ denote the expected profits of the optimal strategies S_1 and S_2 associated with γ_0 and γ_1 , respectively. Then because of the linearity of the relationships, the minimum expected loss for failing to employ the optimal strategy must be the minimum of

$$\epsilon(S_1, \gamma_0) - \epsilon(S_2, \gamma_0)$$

and

$$\epsilon(S_2, \gamma_1) - \epsilon(S_1, \gamma_1),$$

	s = 1	s = 2
F=1	1 - γ	$1-\gamma$
$F = 2 \dots$	γ	γ

TABLE 3
Conditional Distributions of the Amount of the Forecast Given the Demand

where, of course, it is implied that each of these quantities is nonnegative. To choose a strategy which minimizes the expected loss for failing to employ the optimal strategy one would adopt the strategy that is associated with the subtrahend of the above minimum difference. For example, if $\epsilon(S_2, \gamma_1) - \epsilon(S_1, \gamma_1)$ is the minimum expected loss, then one would opt the strategy S_1 .

Properties of the Optimal Strategies

Let us define a *single-act* strategy as one consisting of ordering the same quantity x regardless of the amount of the forecast while a *mixed-act* strategy denotes one wherein the quantities ordered for each amount forecast are not all equal. Then it is clear that for any single-act strategy, the associated $\epsilon(S)$ is independent of γ and hence its graph appears as a straight line parallel to the γ -axis.

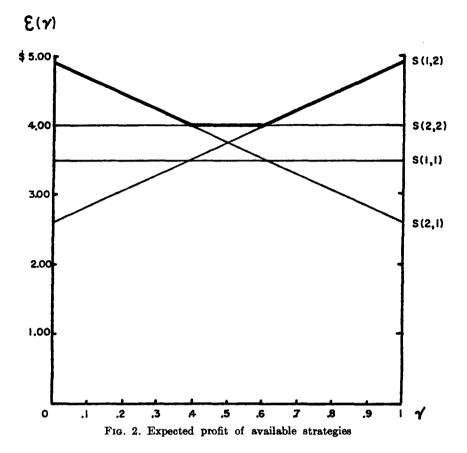
A number of insights may now be discerned. First, under these conditions Figure 1 reveals that, for any γ , there exists a mixed-act strategy which is superior to any single-act strategy. And since the best single-act strategy is the optimal strategy for the non-forecasting case it follows that, for this example, one is always better off in forecasting and employing the corresponding optimal mixed-act strategy than in ignoring the forecast. This proposition holds for any γ .

The second important feature of Figure 1 is that the locus of expected profits for the optimal strategies is not a monotonically increasing function of γ . Beginning at $\gamma = 0$ the optimal $\epsilon(S)$ or, more informatively $\epsilon(\gamma)$, declines with γ until it reaches a minimum at $\gamma = .286$. Beyond this point, however, $\epsilon(\gamma)$ increases with γ , reaching its previous high of \$6.10 when $\gamma = .64$, and then increasing until it reaches its absolute maximum of \$7.00 when $\gamma = 1$.

At first sight, the phenomenon of $\epsilon(\gamma)$ increasing as γ decreases from .286 is puzzling. The explanation is immediately at hand. The forecast with $\gamma=0$, in that it is always wrong, is in fact a "better" forecast than one with $0<\gamma\leq .64$. Starting from $\gamma=.286$, the more inaccurate the forecast, however, the more profitable becomes the adoption of a strategy which negates the forecast. This point can be clarified if we examine a two-event example.

A Two-Event Example

Suppose the probabilities associated with the events z=1 and z=2 are .6 and .4 respectively and that the profit matrix associated with this problem is obtained by deleting row 3 and column 3 of Table 1. Assume that Table 3 contains the conditional distributions of F given z.



The expected profit function for each of the four possible strategies is plotted in Figure 2. We see that for $\gamma \leq (.9/2.3)$, the strategy which maximizes $\epsilon(\gamma)$ is S(2, 1) while for $(.9/2.3) < \gamma < (1 - .9/2.3)$ the best strategy is S(1, 1), a single-act strategy. Finally, if $\gamma \geq (1 - .9/2.3)$ the optimal strategy is S(1, 2).

In this case we note that the locus of the expected profits of the Bayes strategies is symmetrical about $\gamma=.5$. As γ decreases from .9/2.3, it becomes more profitable to adopt a strategy which negates the inaccurate forecast. Hence, we choose a strategy which denies the forecast, which, in turn, has an expected profit equivalent to the optimal expected profit of a more accurate forecast. Therefore, under these conditions, a completely inaccurate forecast, i.e., $\gamma=0$, becomes equivalent to a perfect forecast.

The Worth of an Imperfect Forecast

Figure 2 reveals that, in the absence of a forecast, the best single-act strategy, i.e., the optimal strategy for the non-forecasting case, is S(2, 2) with an expected profit of \$4.00. An important feature of Figure 2 is the fact that for every value of γ , the expected profit of a forecast is equal to or greater than the expected

profit of uncertainty, \$4.00. In other words, under these conditions if the cost of forecasting is zero, then there is nothing lost in attempting to forecast. If the cost of the forecast is not zero, then the decision to forecast depends upon both the cost of the forecast and γ . The maximum amount we would expend on a forecast, however, must be the difference between the expected profit of a forecast and the expected profit of uncertainty or $\epsilon(\gamma)$ — \$4.00.

When to Improve a Forecasting Procedure

By an extension of the previous analysis it is possible to develop insights on the conditions under which attempts to increase γ are warranted.

Referring again to our three-event illustration, suppose we were confronted with a forecasting procedure in which $\gamma = .1$. If someone were to present us with a new procedure with double the accuracy of our current method, i.e., with a procedure in which $\gamma = .2$, it is clear that the optimal use of this new procedure would yield an expected profit less than the quantity we are generating with our less accurate procedure, $\gamma = .1$. Since $\epsilon(.2) < \epsilon(.1)$, we would not make an offer for this new procedure, and, in addition, if it were offered to us gratuitously it is clear that we would reject it and employ our less accurate procedure.

It is apparent that the new procedure must provide $\epsilon(\gamma) = -1.2(.1) + 6.10$ or \$5.98 in order to induce us to accept it. Suppose, in fact, that γ for the proposed procedure is .9. To determine how much to pay for the procedure, as opposed to a given forecast for given γ , we first determine the expected gain of the forecasting procedure. This is equal to $\epsilon(.9)$ for the proposed procedure less $\epsilon(.1)$, the optimal expected profit of our existing procedure. This yields 2.5(.9) + 4.50 or \$6.75 which, when \$5.98 is subtracted, leaves \$.77 per period, an income stream which may be capitalized at the firm's appropriate rate of discount. The present value so obtained must represent the maximum outlay for the proposed forecasting procedure.

The General Case

The characterization of forecasts with a common probability of being correct for any given demand (or event) is not necessary. Turning to the case where this condition does not hold, compact expressions for evaluating the strategies, can still be derived.

Let there exist a payoff matrix, as illustrated below, which associates a payoff a_{ij} for each combination of events $E_j(j=1,2,\cdots,n)$ and acts $A_i(i=1,2,\cdots,n)$. Let there exist, for each E_i , a conditional

Payoff Matrix Events E

		1	2 · ·	$\cdot n$.
	1	a 11	a ₁₂ ···	· a1*
	2	a21	a22 ···	· a28
Acts A	:	a ₁₁ a ₂₁ :	:	:
	_	۱	a	

probability distribution of the forecast F where the probability of

$$F_k(k=1,2,\cdots,n),$$

given E_i , is denoted by P_{ki} . We have previously labeled such an array a forecast matrix, thus

where $\sum_{k=1}^{n} p_{ki} = 1$ for all i.

Since, to each of the n values of F, there are n distinct acts from which one may choose, it follows that there are n^n different strategies corresponding to the problem at hand. When n is even moderately large, direct enumeration of the possible strategies becomes costly. As is well known from standard references (Feller [2, pp. 104-14], and Chernoff and Moses [1, pp. 166-83]), computational economies can be introduced by working with the a posteriori probability distribution of demand.

Let $A_{m(k)}$ denote the optimal act in the face of the forecast F_k and let

$$\epsilon(A_{m(k)} \mid F_k) = \max_j \epsilon(A_j \mid F_k)$$

where

$$\epsilon(A_j \mid F_k) = \sum_{i=1}^n a_{ji} P(E_i \mid F_k),$$

and where $P(E_i \mid F_k)$ denotes the conditional probability of E_i given F_k . Then

$$\epsilon(A_j \mid F_k) = \sum_{i=1}^n a_{ji}(p_{ki} d_i/p_k d)$$

where $p_k = (p_{k1}, p_{k2}, \dots, p_{kn})$ is the kth row of the forecast matrix, and $d = (d_1, d_2, \dots, d_n)$ is a column vector of elements which represent the probability of E_i . The optimal strategy for a given forecast matrix may then be written as $S(A_{m(1)}, A_{m(2)}, \dots, A_{m(n)})$, while the optimal expected profit associated with the forecast matrix is

$$\epsilon(S) = \sum_{k=1}^{n} \epsilon(A_{m(k)1} \mid F_k) p_k d.$$

Observe that this procedure constructs the optimal strategy by calculating n^2 expected values rather than n^n which are required by the direct method.

As before, the expected gain of an imperfect forecast is obtained by subtracting the expected profit of the optimal strategy for the non-forecasting case from the expected profit of the optimal strategy for the given forecast matrix. The expected gain of a forecasting procedure is likewise secured by deducting the expected profit of the efficient strategy associated with the existing forecast matrix from the expected profit of the efficient strategy corresponding to the proposed forecast matrix.

Under certain conditions, however, the expected profit of each optimal strategy for the multiple event case may be obtained, as in Figure 1, as a series of connected line segments. For the set of $\epsilon(\gamma)$ to be a piece-wise linear function of γ , it is sufficient that the probability of each forecast, given z, be expressed as a linear function of γ .

The most restrictive aspect of this requirement is that γ must be the same regardless of the level of demand. In many instances this may be approximately the case and therefore the condition is tolerably satisfied. The second requirement, namely, that the probability of a correct forecast be expressed as a linear function of γ , can always be satisfied since, given γ , any probability can be expressed as proportional to γ .

Other Remarks

While our examples were drawn from the area of elementary inventory theory, we want to stress the complete generality of the analysis developed in this paper. It is relevant to any decision involving the use, potential or actual, of an imperfect forecast where maximization of an expected value is a reasonable criterion. This includes, for example, decisions which are based on forecasts of security prices, interest rates, sales of a product, amount of cash in flow, or the excess reserves of a commercial bank. In the area of fiscal and monetary policy it would include, for example, decisions predicated on forecasts of the GNP, the money supply, or aggregate excess reserves of the banking system.

Data for implementing the above analysis is often at hand. For example, the authors have personal knowledge of several firms that have, in a number of decision processes and for many periods, kept records of forecasts and the corresponding events. When these historical frequencies are smoothed, they can provide the essential conditional distributions of the forecasts.

Summary

Given the probability distribution of the forecast, we have shown how to obtain an optimal strategy which will consist of an ordering of acts, one corresponding to each possible forecast. If the probability of a correct forecast is approximately constant over all possible events, then the expected profit of an imperfect forecast can be expressed as a price-wise linear function of γ and it is possible to find by inspection the optimal strategies for all γ . The analysis suggests that it is often optimal to forecast, even if grossly inaccurate, rather than not forecast and employ the best single-act strategy from the set of all single-act strategies. When forecasts are available we would employ a single-act strategy only as a special case. Assuming that the cost of an individual forecast is zero, our three-event illustration revealed, for example, that the optimal strategy was always a mixed-act strategy while in the two-event example, a single-act strategy is used only if the probability of a correct forecast were to lie in the least effective interval, namely, $(.9/2.3) < \gamma < (1 - .9/2.3)$.

It is clear that knowledge of the probability distribution of forecasts can yield more efficient decisions. The worth of this information is measured by the ex-

pected gain of a forecast which is the difference between the expected profit of a forecast and the expected profit of the optimal act for the non-forecasting case. The concept of the expected value of perfect information is then a special case of the more general notion, the expected gain of a forecast. If the cost of the forecast is less than the expected gain of the forecast, it is profitable to forecast and employ the corresponding optimal strategy. It follows that if the expected gain of the forecast is zero, at least one of the optimal strategies is a single-act strategy and the forecast will be of no value to the decision maker. If the cost of the forecast is greater than zero, it will not be worthwhile to forecast, while if the cost of the forecast is zero, one would be indifferent between employing the forecast and rejecting it.

If the cost of the improvement is zero, a sufficient condition for attempting to improve a forecasting procedure is that the $\epsilon(\gamma)$ associated with the improved value of γ be at least as large as the $\epsilon(\gamma)$ corresponding to the existing procedure. Moreover, the maximum amount one would spend on the new procedure is measured by the present value of the stream: The expected gain of the forecasting procedure per period, i.e., the difference between $\epsilon(\gamma)$ for the proposed procedure and $\epsilon(\gamma)$ for the existing method, over the lifetime of the procedure. Finally, we note that greater accuracy in forecasting, in the sense of increasing γ , is not always a desirable objective even if it can be obtained at zero cost. A more accurate procedure is adopted if there exists a strategy corresponding to the new forecast matrix with an expected profit greater than the expected profit associated with the optimal strategy for the existing forecast matrix.

References

- CHERNOFF, H. AND MOSES, L. E., Elementary Decision Theory, New York, John Wiley and Sons, Inc., 1959.
- FELLEE, W., An Introduction to Probability Theory and Its Applications, Vol. 1, New York, John Wiley and Sons, Inc., 1957.
- 3. RAIFFA, H. AND SCHLAIFER, R., Applied Statistical Decision Theory, Boston, Massachusetts, Division of Research, Graduate School of Business Administration, Harvard University, 1961.
- 4. SAVAGE, L. J., Foundations of Statistics, New York, John Wiley and Sons, Inc., 1954.
- Schlaffer, R., Probability and Statistics for Business Decisions, New York, McGraw-Hill Book Company, 1959.
- STARR, M. K. AND MILLER, D. W., Inventory Control: Theory and Practice, Englewood Cliffs, New Jersey, Prentice-Hall, Inc., 1962.

Copyright 1966, by INFORMS, all rights reserved. Copyright of Management Science is the property of INFORMS: Institute for Operations Research and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.