ASSESSMENT OF THE OUTPUT OF A STOCHASTIC DECISION MODEL

J. B. HARDAKER and A. G. TANAGO*
University of New England

The concept of stochastic dominance is described and its use is illustrated in relation to the evaluation of the output of a systems simulation model of lucerne haymaking in south-west Spain. Two alternative machinery systems are ranked for various lucerne areas using the criteria of stochastic dominance, and these results are compared with those obtained using mean-variance analysis.

Background

In a discussion of the interpretation of systems simulation output for managerial purposes, Dillon [4] notes that it is generally not valid to draw blanket recommendations for managerial action from such output. In principle, the beliefs and preferences of the individual manager are paramount. But as a practical matter, in an agriculture comprising many family-sized firms, the development of general recommendations for complex stochastic decision problems may be a necessary expedient to spread the costs of analysis. In such circumstances, it is clearly desirable that the recommendations made should be as general as possible in terms of the implied assumptions about decision makers' attitudes.

One method of general assessment of risky decision strategies that has been quite extensively used in agricultural analyses is mean-variance (E-V) analysis [5, pp. 27-31]. By use of E-V analysis the set of available strategies can be partitioned into 'efficient' and 'inefficient' sub-sets. A strategy is E-V efficient if no other (pure or mixed) strategy can be found that gives a greater expected payoff with the same variance or the same expected payoff with a lesser variance. By identifying the E-V efficient strategies the choice problem is simplified, if not always resolved, for risk-averse decision makers. Although the use of E-V analysis in agricultural applications has been defended by Anderson [1], the method has some important theoretical limitations [7]. For E-V analysis to be strictly valid either the distribution of outcomes must be normal, or the decision maker's utility function for the payoffs must be quadratic with risk aversion. Otherwise its use may be justified as an approximation by appeal to the Taylor series expansion [5, p. 25] when derivatives of the utility function beyond the second are small. Empirical evidence indicates that many agricultural systems do not generate outcomes that are normally distributed (e.g. [3]), while introspection suggests that the quadratic function is unlikely to be an appropriate representation of commonly-held risk attitudes, since it implies increasing risk aversion as the magnitudes of the payoffs increase [9].

* The authors are indebted to Jock Anderson and John Dillon for helpful comments. The comments of the anonymous referees for this *Journal* are also gratefully acknowledged. A. G. Tanago's graduate work at the University of New England, of which this note is a part, was made possible by the financial support of the Agencia de Desarrollo Ganadero, Ministerio de Agricultura, Spain.

An alternative method of identifying an efficient sub-set of strategies under risk is provided by stochastic dominance analysis. One strategy stochastically dominates another if it would be preferred by all decision makers whose utility functions conform with certain qualitative conditions to be outlined later. First, second and third-degree stochastic dominance have been defined [6, 10, 12].

First-degree stochastic dominance (FSD): the probability function $f(x)$ dominates the probability function $g(x)$ by FSD if, and only if, $F_1(y) \leqslant G_1(y)$ for all $y \varepsilon [a, b]$, with $F_1(y) < G_1(y)$ for at least one value of y.

Second-degree stochastic dominance (SSD): the probability function $f(x)$ dominates the probability function $g(x)$ by SSD if, and only if, $F_2(y) \leqslant G_2(y)$ for all $y \varepsilon [a, b]$, with $F_2(y) < G_2(y)$ for at least one value of y.

Third-degree stochastic dominance (TSD): the probability function $f(x)$ dominates the probability function $g(x)$ by TSD if, and only if, $F_3(y) \leqslant G_3(y)$ for all $y \varepsilon [a, b]$, $F_2(b) \leqslant G_2(b)$, with $F_3(y) < G_3(y)$ for at least one value of y.

In the above, x is a continuous random variable of outcomes of the risky prospects; the closed interval [a, b] is the sample space of both prospects; $F_n(y) = {}_a\int^y F_{n-1}(x)dx$, $F_o(y) = f(x)$, and similarly for $G_n(y)$. Corresponding definitions can be derived for the case where x is a discrete random variable.

Each of these dominance criteria divides the set of possible strategies into efficient and inefficient sub-sets, where a strategy is efficient if, and only if, it is not dominated by another strategy. FSD implies SSD which implies TSD, so that more strategies can be ordered by TSD than by SSD which in turn can order more strategies than FSD.

The use of stochastic dominance is not restricted to distributions of any particular form. However, the existence of a utility function for the decision maker that conforms with the usual axioms of Benoullian theory [5] is assumed. Moreover, the criteria depend upon certain assumptions about the form of this utility function. FSD requires only that utility shall be a monotonically increasing function of the payoff, i.e. $U_1(y) > O$, $y \varepsilon [a, b]$; SSD requires also that the function shall be everywhere risk averse, i.e. $U_1(y) > O$ and $U_2(y) < O$, $y \varepsilon [a, b]$; and TSD requires the additional condition of decreasing risk aversion, i.e. $U_1(y) > O$, $U_2(y) < O$ and $U_3(y) > O$, $y \varepsilon [a, b]$, where $U_n(y)$ is the n-th derivative of the decision maker's utility function.

The formal proofs of FSD and SSD are provided by Hadar and Russell [6], and Whitmore [12] gives a proof of TSD. In this note we confine ourselves to an informal interpretation of FSD and SSD. (For TSD we find an intuitive interpretation inappropriate and must rely on reference to the formal proof.)

The definition of FSD means that the cumulative density function (CDF) of the preferred prospect must lie in part to the right and nowhere to the left of the CDF for the dominated prospect, as illustrated in Figure 1. When the CDFs show one or more intersections, as illustrated in Figure 2, the condition for FSD is not satisfied and the distributions must be examined in relation to the criteria of SSD or TSD.

For a simple illustration of FSD, consider two risky prospects s_1 and

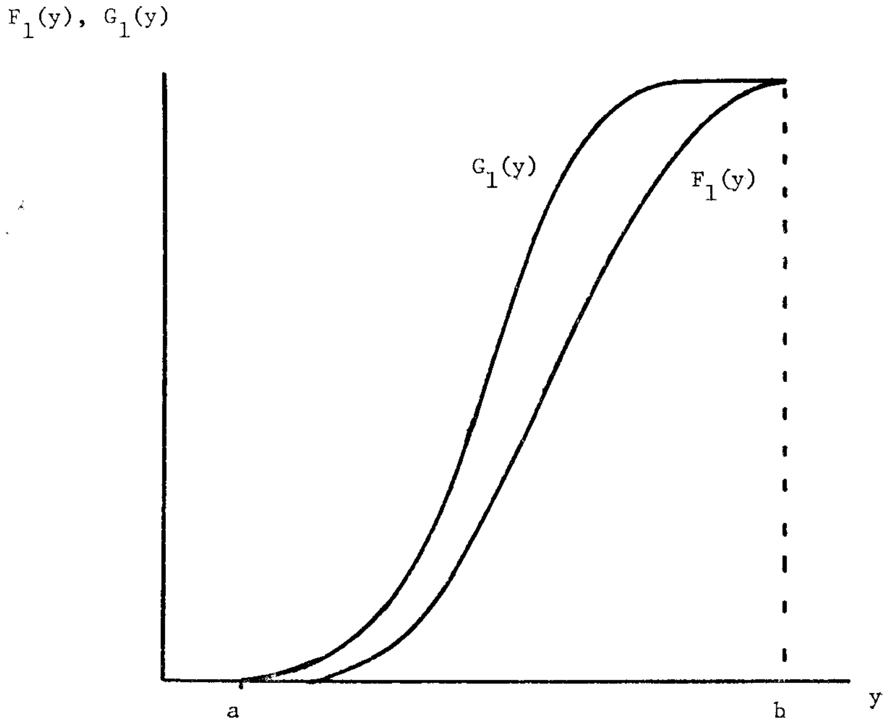

FIGURE 1—First-degree stochastic dominance.

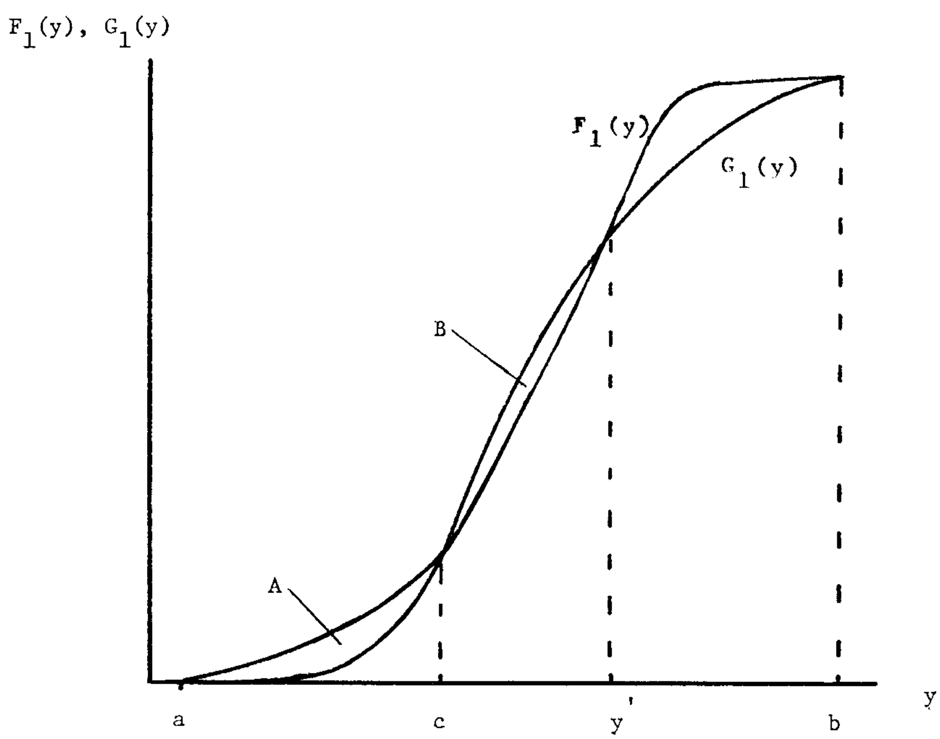

FIGURE 2—No first-degree stochastic dominance.

s_2, each with uncertain payoffs of a and b, $a < b$. Assume the existence of a monotonically increasing utility function u such that $u(a) < u(b)$. Suppose the probability of payoff a for prospect s_1 is p and for s_2 is q. Then the utilities of the two prospects are given by their respective expectations:

$$u(s_1) = pu(a) + (1-p)u(b)$$
$$u(s_2) = qu(a) + (1-q)u(b)$$

The rule for FSD in this simple situation is the self-evident condition that s_1 will be preferred to s_2 if, and only if, $p < q$. That is, all decision makers who prefer more payoff to less will prefer the prospect for which the probability of receiving the lesser payoff is lowest.

Consider now an extension of the above situation where we have not two but n possible payoffs for each prospect. Denote these payoffs, arranged in increasing order of magnitude, by $x_1, x_2, \ldots, x_n$. Under the same assumption of a monotonically increasing utility function, u, we have $u(x_1) < u(x_2) < \ldots < u(x_n)$, i.e. $u(x_i) < u(x_j)$ if, and only if, $i < j$. Let the respective probabilities for s_1 be $p_1, p_2, \ldots, p_n$ and for s_2 be $q_1, q_2, \ldots, q_n$. We can now deduce from the simple example above that s_1 will be preferred to s_2 if $p_j < q_j$ and $p_k > q_k$ where $k > j$ with $p_i = q_i$ for all $i \neq j,k$. That is s_1 will be preferred if it bears a lower probability than s_2 of one less favourable payoff, with a correspondingly increased probability of one more favourable payoff. This is a particular example of the general condition for FSD, which for this discrete case can be written as

$$\sum_{i=1}^{r} p_i \leqslant \sum_{i=1}^{r} q_i \text{ for } r = 1, 2, \ldots, n$$

with strict inequality for at least one value of r. This condition implies a consistent weighting of probabilities towards the more favourable payoffs for s_1 compared with s_2.

For an interpretation of SSD it is convenient to rewrite the condition as

$$\int_a^y F_1(x)dx - \int_a^y G_1(x)dx \leqslant 0.$$

In applying this rule to the situation shown in Figure 2, and taking the particular value of $y = y'$, illustrated, we can divide the interval $[a\ y']$ into two at c where the curves intersect. The condition for SSD can now be expanded to give

$$\left[\int_a^c F_1(x)dx - \int_a^c G_1(x)dx\right] + \left[\int_c^{y'} F_1(x)dx - \int_c^{y'} G_1(x)dx\right] \leqslant 0.$$

The term in the first set of square brackets is the negative of area A in the figure, while the term in the second set of square brackets is the area B. The condition for SSD is least likely to be satisfied for the value of y' shown, and we have therefore deduced that the probability function f(x) will dominate the function g(x) by SSD if the area A is not less than the area B.

For an explanation of this result we need to recall that SSD requires the assumption of diminishing marginal utility ($U_2(y) < 0$). Assume prospect s_1 has payoffs distributed according to f(x) and prospect s_2 has payoffs distributed according to g(x). Then the utility gain for s_1 over s_2 associated with the reduced probability of low payoffs represented by area A must be less than the utility loss associated with the higher probability of intermediate outcomes represented by area B, since area A $\geqslant$ area B and the marginal utility of x is greater in the interval [a c] than in the interval [c y′].

Application

Both E-V analysis and stochastic dominance analysis were applied in the assessment of the results of a systems simulation model of lucerne haymaking on irrigated farms in south-west Spain. The objective of the study was to choose between two alternative sets of haymaking equipment. The system was stochastic, being affected by random weather variables. The principal output variable of interest was gross margin, which was measured in replicated experimental treatments for each machinery system on a number of representative farm areas. More complete details of the model are provided elsewhere [11].

For E-V analysis the first two moments of the distribution of outcomes were estimated directly from the experimental results for each treatment. Since only two strategies were considered for each farm area, any E-V ordering resulted in the identification of a single E-V efficient strategy. E-V analysis proved capable of ordering the machinery systems for five of the ten farm areas considered. However, inspection of the sampled outcomes indicated that the payoffs were not normally distributed, so that strictly valid use of the E-V criterion depended upon the assumption of quadratic utility functions.

Application of stochastic dominance analysis involved the steps described below.

(1) The cumulative density function (CDF) for each experimental treatment was estimated from the set of outcomes obtained from the model using a procedure for smoothing sample data [2].

(2) The CDFs for both machinery systems were plotted on the same graph for each farm area considered. Cases of FSD could then be identified by visual inspection. For example, Figure 3, which corresponds to an area of 30 ha, shows that the values of the CDF for machinery system 2 are everywhere larger than those of machinery system 1. Thus system 1 is stochastically larger than system 2 and dominates in the sense of FSD. Where FSD did not occur, as illustrated in Figure 4, which relates to a lucerne area of 40 ha, further analysis involving SSD or TSD was needed. In fact, the mechanization strategies could be ranked by FSD in all cases but one.[1]

(3) The analysis of SSD was accomplished by dividing the range of x into a large number of discrete intervals of width $\triangle x$. The area under the CDF curve was then estimated as the sum of the series of rectangles of width $\triangle x$ and of height equal to the ordinate of the midpoint of each interval.[2] i.e.

[1] The reason for this favourable result was that the shapes of the distributions were similar for both systems but as the harvest area was varied, the relative average fixed costs were affected. In terms of the FSD analysis, this resulted in shifts in the relative positions of the two CDFs with clear dominance for system 1 on the smaller areas and for system 2 for larger areas.

[2] This approximate method was used since algebraic functions for the CDFs were not fitted. Had such algebraic functions been obtained, direct integration might have been possible.

$$F_2(y_n) \doteq \Delta x \sum_{x=1}^{n} F_2(x_i - \tfrac{1}{2}\Delta x),$$

where $x_i = x_{i-1} + \Delta x$;
$n = (y_n - x_o)/\Delta x$;
$y_n \varepsilon [x_o, x_1, \ldots, x_m]$;
$x_o = a$;
$x_m = b$.

The estimates $F_2(y)$ and $G_2(y)$ were compared for each value of y for assessment of SSD. In the case studied, SSD was not found, so further analysis using TSD was needed.

(4) The analysis of TSD was accomplished by estimating the area under the relevant curves using the same method as for SSD. i.e.

$$F_3(y_n) \doteq \Delta x \sum_{i=1}^{n} F_2(x_i - \tfrac{1}{2}\Delta x),$$

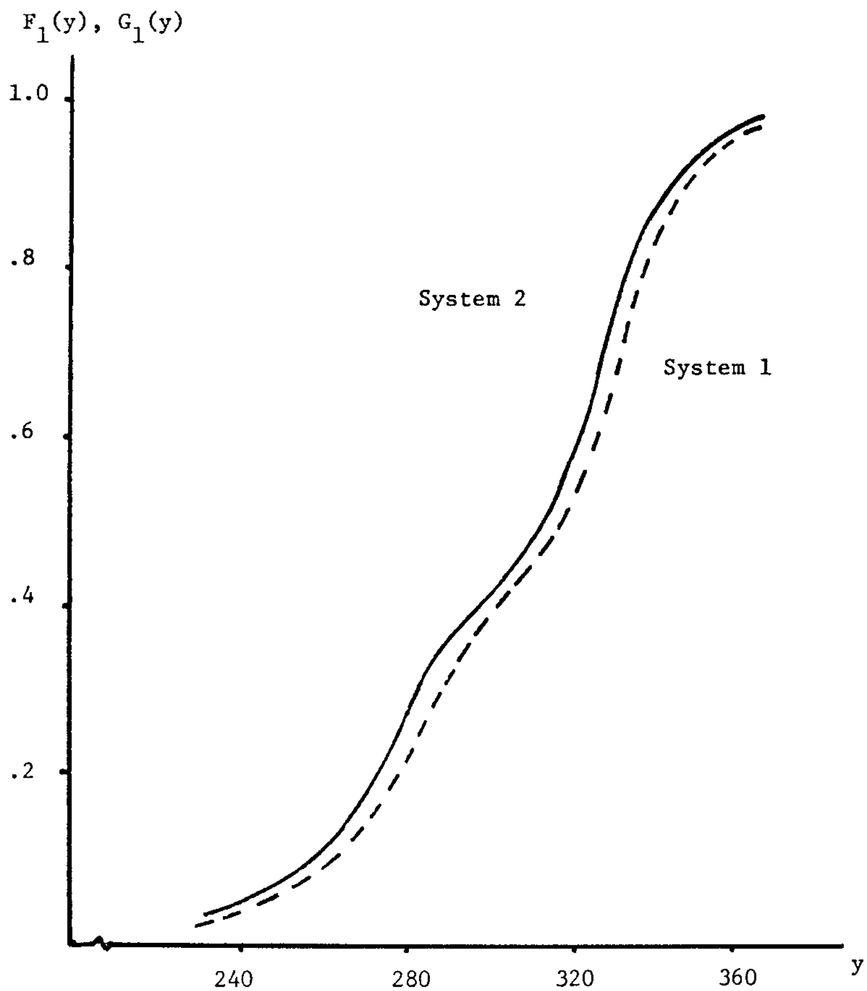

FIGURE 3—Cumulative density functions fitted to assessed fractiles (area of 30 ha).

where terms are defined as before. The results of this analysis enabled the decision strategies for the remaining case to be ranked since it was found that for a lucerne area of 40 ha, system 1 dominated system 2 by TSD.

Discussion

A comparison of the results obtained from E-V and stochastic dominance analysis is presented in Table 1. It can be seen that, in this application, stochastic dominance criteria were able to identify the optimal strategy for all ten cases considered, whereas E-V analysis could rank the strategies in only half the cases. The generality of this apparent superiority of stochastic dominance over E-V analysis must be suspect since it is somewhat at variance with the results of Porter and Gaumnitz [8] who concluded, on the basis of a more comprehensive analysis than ours, that except for highly risk-averse decision makers, the choice between the two methods of analysis was not critical.

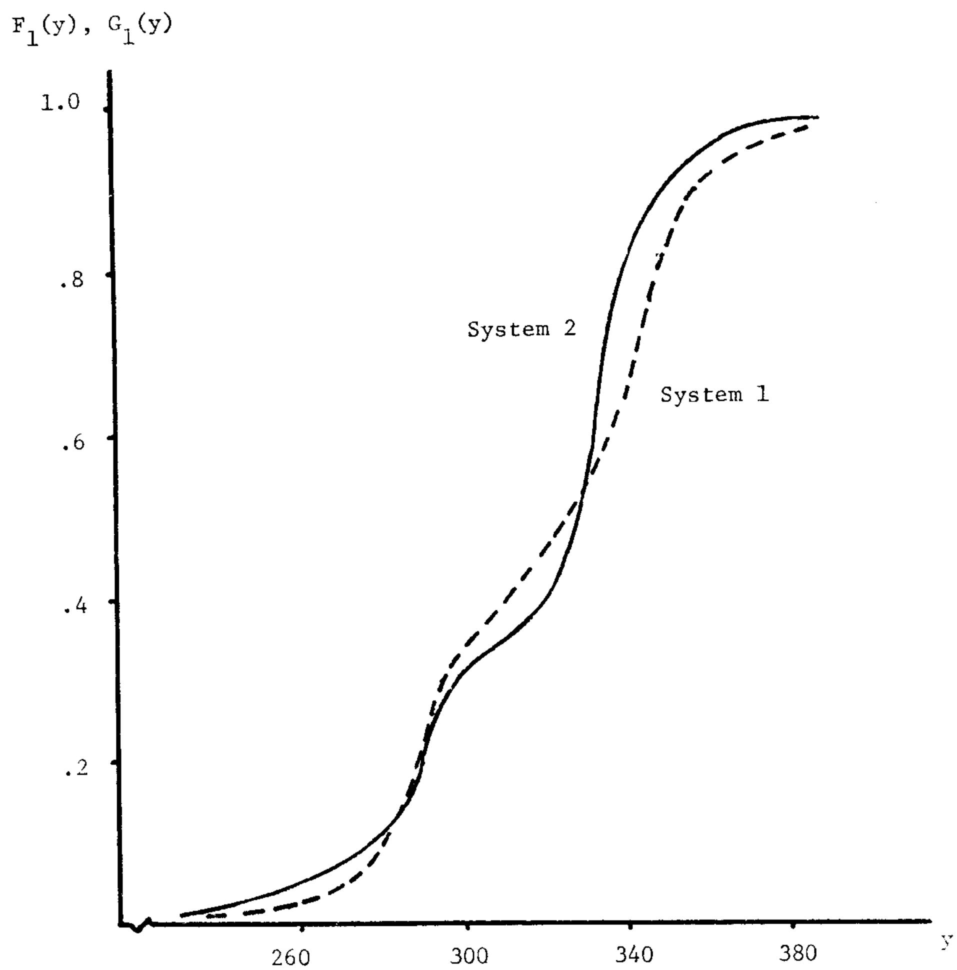

FIGURE 4—Cumulative density functions fitted to assessed fractiles (area of 40 ha).

TABLE 1

Efficient Strategies for Each Area Identified Using Different Methods of Analysis[a]

Area (ha)	E-V Analysis	Stochastic Dominance FSD	SSD	TSD
10	1	1	1	1
20	1	1	1	1
30	1	1	1	1
40	0	0	0	1
50	0	2	2	2
60	2	2	2	2
80	2	2	2	2
100	0	2	2	2
120	0	2	2	2
150	0	2	2	2

[a] The numeral 1 indicates that machinery system 1 dominates system 2, while the numeral 2 indicates that the reverse applies. Zero is used when the criterion does not lead to the selection of a unique efficient strategy.

An important difference between the two methods is the ability of stochastic dominance analysis to eliminate low return, low variance, E-V efficient strategies [8, p. 445]. On the other hand, as Whitmore notes [12, p. 458], there are many practical situations where preference between two risky prospects cannot be established by stochastic dominance analysis. For example, prospect s_1 may appear attractive relative to prospect s_2 except that there is a small, but non-zero probability of payoffs for s_1 more adverse than the worst payoff for s_2. Regardless of the rest of the distribution of the payoffs of the two prospects, stochastic dominance criteria will never establish that s_1 is preferred to s_2. In such cases E-V analysis may provide a reliable ranking of the two prospects.

In this application stochastic dominance analysis has been shown to advantage partly because only two decision options were considered. In other applications with more than two options it is to be expected that the stochastically dominant set will contain more than one strategy. Identification of the single optimal strategy will then require either a subjective appraisal by each individual decision maker, or more complete specification of each decision maker's utility function. Note too that the valid use of stochastic dominance analysis, like E-V analysis, requires the distributions of payoffs to conform with the decision maker's subjective beliefs. Such a requirement is likely to be satisfied only when the decision model has been developed using the individual decision maker's subjective probabilities, or when the decision system, including the stochastic components affecting it, is well understood, so that the distributions of outcomes can be regarded as 'public'.

Although not presenting any severe operational difficulties, stochastic dominance analysis is perhaps not quite so simple to apply as E-V analysis. Certainly, there is no obvious way in which the criteria can be incorporated into a mathematical programming model, as is possible with E-V analysis. However, the chief merit of stochastic dominance over E-V analysis lies in the greater generality of the underlying assumptions. The theoretical superiority of the technique and its relatively straight-forward application, indicate that it deserves more attention as an analytical approach to decision making under risk.

References

[1] Anderson, J. R., 'Risk Aversion and Polynomial Preference', *Australian Economic Papers,* in press, 1973.

[2] Anderson, J. R., 'Sparse Data, Climatic Variability and Yield Uncertainty in Response Analysis', *American Journal of Agricultural Economics,* in press, 1973.

[3] Day, R. H., 'Probability Distributions of Field Crop Yields', *Journal of Farm Economics,* 47 (3): 713-741, August 1965.

[4] Dillon, J. L., 'Interpreting Systems Simulation Output for Managerial Decision-making', in J. B. Dent and J. R. Anderson (Eds), *Systems Analysis in Agricultural Management,* pp. 85-122, Wiley, Sydney, 1971.

[5] Dillon, J. L., 'An Expository Review of Bernoullian Decision Theory in Agriculture: Is Utility Futility?', *Review of Marketing and Agricultural Economics,* 39 (1): 3-80, March 1971.

[6] Hadar, J., and W. R. Russell, 'Rules for Ordering Uncertain Prospects', *American Economic Review,* 59 (1): 25-34, March 1969.

[7] Hanoch, G., and H. Levy, 'Efficient Portfolio Selection with Quadratic and Cubic Utility', *Journal of Business,* 43 (2): 181-189, April 1970.

[8] Porter, B. R., and J. E. Gaumnitz, 'Stochastic Dominance vs. Mean-variance Portfolio Analysis: An Empirical Evaluation', *American Economic Review,* 62 (3): 438-446, June 1972.

[9] Pratt, J. W., 'Risk Aversion in the Small and in the Large', *Econometric,* 32 (1-2): 122-136, January-April 1964.

[10] Quirk, J. P., and R. Saposnik, 'Admissibility and Measurable Utility Functions', *Review of Economic Studies,* 29 (2): 140-146, February 1962.

[11] Tanago, A. G., 'Hay-making Machinery Selection Under Risk: A Simulation Approach', unpublished Dip. Agric. Econ. dissertation, University of New England, 1973.

[12] Whitmore, G. A., 'Third-Degree Stochastic Dominance', *American Economic Review,* 60 (3): 457-459, June 1970.