

A GENERALIZED STOCHASTIC DOMINANCE PROGRAM FOR THE IBM PC

Siew Goh, Chao-Chyuan Shih, Mark J. Cochran, and Rob Raskin

Abstract

A microcomputer program to perform Generalized Stochastic Dominance (GSD), Quasi-Second Degree Dominance (SSD), and Quasi-First Degree Stochastic Dominance (FSD) is described. The program is designed to run on IBM-compatible personal computers with a Hercules or CGA graphics adapter. It is menu-driven and has options for GSD, quasi-FSD, quasi-SSD, graphics, and calculations of premiums associated with use of dominant distributions.

Key words: IBM-compatible personal computers, stochastic dominance, applied-risk analysis, expected utility.

Generalized Stochastic Dominance (GSD) has become a popular technique for ordering risky strategies when it is desired to consider more than the first two moments of the underlying probability distributions (Meyer). GSD has become a research tool that is commonly used in applied risk analysis and has been employed in a variety of settings, such as machinery selection (Danok et al.), pest management (Greene et al.; Cochran et al.; Zacharias and Grube), public policy (Kramer and Pope; King and Oamek; Lemieux et al.), crop insurance (Zering et al.), irrigation scheduling (Bosch and Eidman), marketing (Anaman and Boggess; Holt and Brandt; King and Lybecker; Rister et al.), and life insurance (Tauer). Stochastic dominance procedures reduce a choice set of alternative management strategies down to a smaller subset which should include the strategy that maximizes expected utility for the class of relevant risk preferences. The subset is referred to as the efficient set, and its members are risk-efficient (or dominant) strategies for the pref-

erences analyzed.

The GSD program to be described has been written for use on IBM-compatible personal computers in FORTRAN, PASCAL, and ASSEMBLY languages. It uses IBM DOS 3.1, a color graphics (CGA) or Hercules graphics adapter, and about 180K exclusive of data files. Users on machines without either the CGA or the Hercules graphics adapters can still operate the program but should avoid selecting the graphics option.

The program is menu-driven and includes options for quasi-first- and second-degree stochastic dominance (to be described below) and generalized stochastic dominance. It can analyze choice sets of up to 50 distributions of 100 elements each (another version of the program is available which handles choice sets of 100 distributions with 200 elements each, but it requires additional memory). It also has graphics capabilities for displaying up to five of the distributions at one time and can calculate the value of information contained in risk-efficient strategies. This program originated in a mainframe program described by King and Robison.

GENERALIZED STOCHASTIC DOMINANCE

The basis for GSD is the expected utility hypothesis (EUH), which states that strategy F is preferred to strategy G when the expected utility of F exceeds that of G. Strategy G can then be excluded as risk inefficient. Rather than facing the task of measuring exactly the risk preferences of decision makers, the analyst can imprecisely represent them with an interval which should include the relevant preferences. The preference interval is defined as a class of admissible utility functions. In practice, this definition is implemented by the specifica-

Siew Goh is a Research Associate, Chao-Chyuan Shih is a former Research Associate, Mark J. Cochran is an Associate Professor, Department of Agricultural Economics, University of Arkansas; and Rob Raskin is a Graduate Assistant, Department of Atmospheric and Oceanic Science, University of Michigan.

Copyright 1989, Southern Agricultural Economics Association.

tion of bounds on the Pratt-Arrow absolute risk aversion function, $R = -U''(x)/U'(x)$. One distribution is said to dominate another if its expected utility exceeds that of the other for every utility function belonging to that class. By specifying a relatively large class of admissible utility functions, the analyst can minimize the chance of incorrectly excluding the preferred option (a Type I error). However, this is accomplished at the expense of weaker discriminating capabilities (Type II error) as the resulting efficiency sets are typically quite large. The specification of a relatively small class of utility functions produces the converse effects: stronger discriminating capabilities are obtained but with an increased likelihood of a Type I error.

The test for dominance is carried out as follows. By the expected utility hypothesis, distribution F has greater expected utility than distribution G when

$$(1) \int_{-\infty}^{\infty} (f(x) - g(x))U(x)dx > 0,$$

where f and g are the corresponding probability density functions of the cumulative distribution functions F and G, respectively. When this expression is integrated by parts, it becomes

$$(2) \int_{-\infty}^{\infty} [G(x) - F(x)] * U'(x)dx > 0.$$

That is, F is preferred to G if the area under the difference of the cumulative distribution functions, weighted by the marginal utility at each point, is positive. In terms of $r(x)$, if

$$(3) U'(x) = \exp\left(\int_0^x -r(x)\right),$$

then dominance is found when

$$(4) \int_{-\infty}^{\infty} [(G(x) - F(x))] * \exp\left(\int_0^x -r(x)\right)dx > 0.$$

For F to dominate G with respect to a function, the above integral must be positive for all decision makers whose $r(x)$ is within the specified bounds on r . Using optimal control techniques, a limiting $r(x)$ is found which minimizes the above integral over all permissible values of $r(x)$. If this minimum value is positive, the value of the integral will be positive for all other risk preference functions within the stated bounds as well. In such a case, stochastic dominance has been found. If stochastic dominance is not found, the roles of distribution F and distribution G are exchanged in an attempt to identify dominance of G over F. A review of the literature of stochastic dominance can be found in Cochran.

bounds on r , first-degree or second-degree stochastic dominance is often used. First-degree stochastic dominance (FSD) analysis simply places no restrictions on the value of r (although the assumption of positive marginal utility is retained). This methodology eliminates the possibility of a Type I error entirely (assuming that decision makers always prefer more to less and that the probability distributions have been measured without error). Under this criterion, only when one CDF curve never crosses another CDF curve can the inferior distribution be eliminated from the efficient set.

If the restriction is made to include only risk-averse decision makers, a second-degree stochastic dominance (SSD) analysis is performed. As risk aversion is characterized by a concave utility function ($u'' < 0$), r must be positive. Under this criterion, only when the "cumulative of the cumulative" of one distribution curve never crosses that of another can second-degree stochastic dominance occur.

Very often, a distribution may be inferior to another but will not be eliminated solely because its cumulative (under FSD) or its cumulative of the cumulative curve (under SSD) crossed that of another at a single point (the latter case is sometimes referred to as the "lefthand tail" problem). This problem arises when preferences approximating the maxi-min decision rule are included in the preference interval. The quasi-first- and quasi-second-degree stochastic dominance options available in this program are not true applications of these techniques, but with most choice sets the efficient sets produced should be identical. The quasi-first- and quasi-second-degree stochastic dominance options are special cases of GSD. For quasi-first-degree stochastic dominance, the bounds on $r(x)$ are set wide enough to include essentially all observed risk preference behavior (rather than allowed to rise to positive or negative infinity). A similar procedure is used for the quasi-second-degree stochastic dominance as well, but the lower bound is set equal to zero.

The actual upper bound for the two options was set so that the relative risk-aversion coefficient never exceeded 100. The relative risk-aversion coefficient is defined by $rr = r * x$ and represents the elasticity of the marginal utility function. For the quasi-first-degree analysis, the lower bound was set at the negative value of the upper bound.

FIRST-DEGREE AND SECOND-DEGREE STOCHASTIC DOMINANCE

In the absence of information regarding the

PROGRAM OPTIONS

The program is menu driven and contains nine different options. Most are self-explanatory,

but two may require additional information for analysts unfamiliar with the stochastic dominance literature.

INPUT REQUIREMENTS & OUTPUT OPTIONS

The inputs required to run this program consist of: 1) the set of distribution points for each strategy to be evaluated. (It is assumed that each point within a given distribution occurs with equal likelihood.) The distributions can be entered directly or read from a file. Once the distribution points have been entered, they may be stored on floppy or hard disk for subsequent runs; and 2) for generalized stochastic dominance, a pair of (constant) values of R, representing the lower bound and upper bound on admissible preferences.

The program can output any of the following (see Figure 1 for the main menu):

A) the names of the strategies making up the efficiency set for quasi-first-degree, quasi-second-degree, or generalized stochastic dominance (Options 1, 2, and 3, respectively); or

B) the results of all pairwise tests for generalized stochastic dominance between all distributions having expected values greater than a specified value (Option 4); or

C) for users having either a CGA or a Hercules graphics adapter board installed, a graphical output of up to five of the distributions on one graph (Option 5); or

D) the stochastic value of information from a pairwise comparison between a dominant and user-specified distribution (Option 6); or

E) descriptive statistics on the data set (Option 7).

The names of the efficient strategies (from Options 1, 2, or 3) or the pairwise comparison results (from Option 4) can be stored in a file (on a floppy or hard disk) for later evaluation.

The output from the most commonly used options are displayed. An example from an irrigation scheduling problem is presented to demonstrate most of the options. The data from this problem are displayed in Table 1. The first line in the file indicates that there are eight distributions with 23 elements each. The names of the management strategies associated with the eight distributions are offset beginning in column 1. In this case, outcome variables are expressed in terms of annual net income per acre.

Options 1 and 2 perform the quasi-first- and quasi-second-degree stochastic dominance already described. It should be remembered that these are not true applications and in rare cases may not replicate the actual efficient sets of these criteria. However, in the overwhelming number of cases the results of the FSD and SSD should be consistent with those of the quasi-first- and quasi-second-degree stochastic dominance generated with this program. Differences may arise with choice sets that contain distributions that are extremely similar and/or encounter a left-hand tail problem in the lower end of the outcome range.

Options 3 and 4 perform generalized stochastic dominance (also known as stochastic dominance with respect to a function and the Meyer criterion). The difference between the two

STOCHASTIC DOMINANCE WITH RESPECT TO A FUNCTION

OPTIONS:

- 1) QUASI FIRST DEGREE STOCHASTIC DOMINANCE
- 2) QUASI SECOND DEGREE STOCHASTIC DOMINANCE
- 3) STOCHASTIC DOMINANCE WITH RESPECT TO A FUNCTION
- 4) STOCHASTIC DOMINANCE WITH RESPECT TO A FUNCTION (REPORT RESULTS OF ALL PAIRWISE COMPARISONS (COMPARE ONLY DISTRIBUTIONS WITH EXPECTED VALUES GREATER THAN A SPECIFIED VALUE))
- 5) GRAPH THE DISTRIBUTION DATA
- 6) CALCULATE THE AMOUNT BY WHICH THE DOMINANT DISTRIBUTION CAN BE LOWERED BEFORE IT NO LONGER DOMINATES THE OTHER DISTRIBUTION
- 7) COMPUTE THE MEANS, STANDARD DEVIATION, AND SKEWNESS
- 8) ENTER DOS COMMAND
- 9) EXIT STOCHASTIC DOMINANCE PROGRAM

SELECT OPTION DESIRED BY TYPING APPROPRIATE NUMBER AND <RETURN>

Figure 1. Sample Main Menu Display.

TABLE 1. A SAMPLE DATA SET

8	23							
two		MEAN=		-28.31	S.D.=		60.48	
	16.71	-16.67	-112.02	-41.95	-115.95	-100.13	-26.01	
	-37.36	109.67	-49.64	-44.17	-6.45	-30.06	-1.78	
	19.84	4.77	-72.20	24.93	-64.32	50.34	-166.40	
	19.29	-11.59						
nirrg		MEAN=		-15.22	S.D.=		55.89	
	52.11	15.26	-81.04	-6.90	-83.18	-90.36	-38.85	
	-21.72	120.34	-57.12	-10.40	15.82	-33.50	34.85	
	39.06	36.13	-72.47	35.80	-57.88	-.89	-116.38	
	-26.81	-1.84						
cap50		MEAN=		70.25	S.D.=		27.51	
	74.03	99.45	25.37	58.83	74.69	55.16	73.94	
	89.68	112.82	75.54	93.28	85.33	96.31	85.63	
	41.29	49.49	79.97	91.74	50.40	65.27	-11.72	
	59.09	90.08						
t0530		MEAN=		62.86	S.D.=		26.18	
	59.77	80.04	25.47	49.48	39.97	61.16	68.31	
	85.40	114.44	33.37	72.38	87.65	97.87	58.59	
	36.02	42.73	85.41	95.84	68.64	42.14	8.38	
	50.84	81.86						
t0630		MEAN=		57.11	S.D.=		32.23	
	55.39	74.23	8.33	48.66	38.22	49.01	69.17	
	77.55	111.66	34.04	73.39	85.49	95.87	47.32	
	41.34	34.93	87.55	95.08	58.12	42.92	-39.33	
	50.52	74.08						
t0330		MEAN=		67.08	S.D.=		24.31	
	78.59	90.93	28.75	27.64	61.82	56.20	70.19	
	84.43	115.10	52.26	91.32	81.57	94.84	87.39	
	39.41	43.29	81.43	89.45	63.56	43.14	27.71	
	58.42	75.35						
t0430		MEAN=		62.69	S.D.=		26.07	
	55.48	98.38	25.51	51.04	40.94	45.11	67.49	
	82.14	115.97	31.61	71.33	86.17	96.12	55.02	
	40.31	37.06	86.08	94.01	51.98	56.08	18.12	
	53.33	82.60						
cap60		MEAN=		67.72	S.D.=		19.49	
	72.25	93.53	45.68	47.23	66.97	44.80	71.35	
	86.02	96.30	79.61	81.45	76.07	85.88	89.90	
	35.57	43.73	68.06	81.06	69.02	49.52	30.95	
	57.96	84.65						

TABLE 2. SAMPLE OUTPUT FROM OPTION #3

R1=	.001000	R2=	.020000	DISTRIBUTION FILE: soy987
cap50				
cap60				

options arises from the number of pairwise comparisons that are made. The output of option 3 is only a listing of the efficient set. Whereas

option 4 will generate a table displaying the results of all pairwise comparisons, option 3 will not always perform all of the pairwise compari-

TABLE 3. SAMPLE OUTPUT FROM OPTION #4

R1=	.001000	R2=	.020000	DISTRIBUTION FILE: soy987
two	-0000	000		
nirrg	1-000	000		
cap50	11-11	?1?		
t0530	110-1	0?0		
t0630	1100-	000		
t0330	11?11	-10		
t0430	110?1	0-0		
cap60	11?11	11-		

sons since it will exclude from any further analysis a distribution which is dominated by some other. This option should be used with large choice sets and when the analyst is only concerned with identifying the efficient set. To facilitate the speed with which option 3 will perform the analysis, those distributions which are believed to be most likely to dominate should be entered into the input file first. When the analyst has interest in the complete ranking of the alternative strategies, then option 4 should be selected. It should also be noted that this algorithm uses only constant absolute risk-aversion functions, so preferences which vary with income cannot be accommodated.

Sample outputs from options 3 and 4 appear in Tables 2 and 3. Table 2 exhibits the output

from option 3, consisting of a statement of the preference interval expressed in terms of lower and upper bounds on the Pratt absolute risk-aversion function, the identification of the input file containing the distribution data, and a listing of the strategies which are members of the efficient set. This particular output can be interpreted as follows: strategies Cap50 and Cap60 dominate all other strategies appearing in the choice set from file soy 987, displayed in Table 1. Table 3 displays the output of option 4. It contains the statement of the preference interval and identification of the input file as before, but it features a matrix of results from all pairwise comparisons rather than simply exhibiting the membership in the efficient set. The rows are labeled with the names of the distributions, and

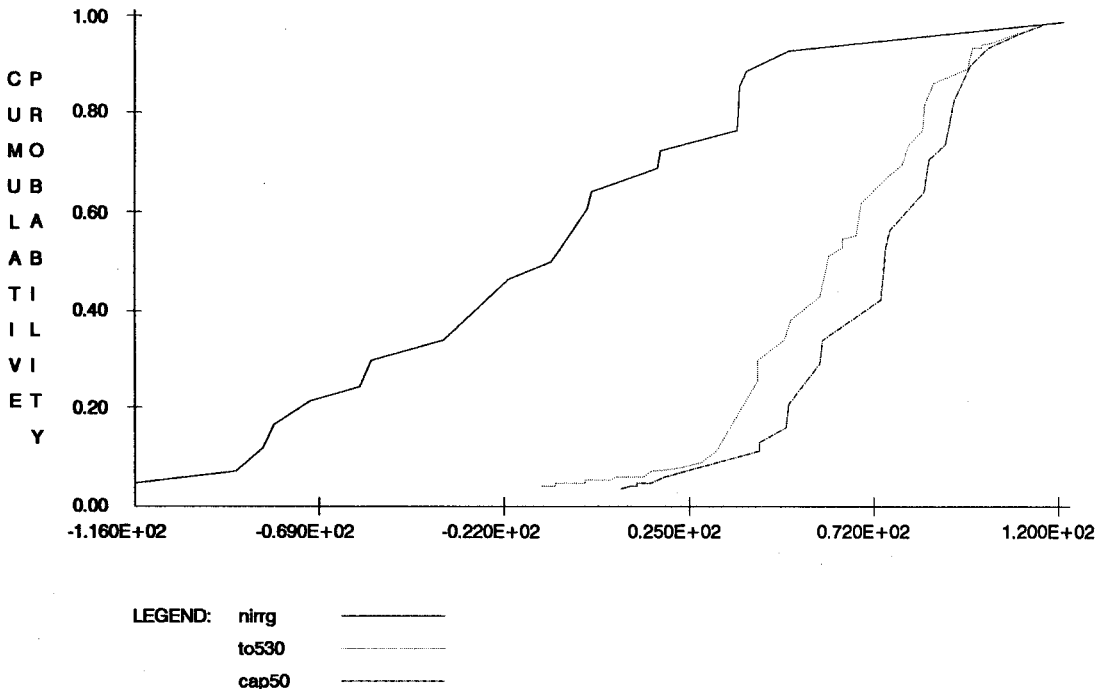


Figure 2. Sample Graph.

TABLE 4. SAMPLE OUTPUT FROM OPTION #6

R1=	.001000	R2=	.020000	DISTRIBUTION FILE: soy987	
THE DOMINANT DISTRIBUTION IS:			cap50		
THE USER-SPECIFIED DECREMENT SIZE IS:			2.000		
DISTRIBUTION NAME	LOWER BOUND	ITERATIONS	UPPER BOUND	ITERATIONS	
t0430	.450000E+01	5	.850000E+01	10	
t0530	.500000E+01	7	.800000E+01	13	
nirrg	.867500E+02	47	.103250E+03	58	

the columns represent the distributions in the same order. The interpretation of the symbols is as follows: 1 indicates that the distribution in the row dominates the distribution of the column; 0 indicates that the distribution in the row is probably dominated by the distribution in the column; and ? indicates that neither distribution dominates for the given class of decision makers represented by the preference interval. The dashes (-) down the diagonal in the center of the matrix indicate that distributions are not compared to themselves. For example, strategy "t0530" dominates strategies "two," "nirrg," and "t0630." It is dominated by strategies "cap50," "t0330," and "cap60." Decision makers represented by the preference interval are indifferent between strategies "t0530" and "t0430." Finally, the efficient set consists of the strategies whose rows contain only "1" or "?," hence in this case, strategies "cap50" and "cap60."

An option to graph up to five different distributions is next. This option is designed for use with machines accessing either a CGA or a Hercules graphics card and should not be selected by users without this capability. The horizontal axis is defined for each graph and is labeled dependent upon the range of outcome values found in the specified distributions. A sample graph, displayed in Figure 2, shows the cumulative distribution functions of three strategies identified in the legend. The graphs can be used to recognize left-hand tail problems and identify distributions of close proximity.

The stochastic value of information contained in risk-efficient management strategies can be identified with option 6. Traditionally, a single estimate of the value of information has been provided (Hilton). Option 6 calculates an upper and lower bound on the value of information as a function of risk preferences. The theoretical foundation and an application appear in Cochran and Mjelde. It is an extension of work by Byerlee and Anderson, Bosch and Eidman, and Rister et al. The option measures the amount that a class of decision makers (defined by the

preference interval) would be willing to pay, in each state of nature, and remain indifferent between a dominant distribution and an inferior alternative. In practice, this translates into identifying the magnitude of a parallel shift in the dominant distribution that would be necessary to eliminate the dominance and produce a change in the efficient set. The estimated value reflects the most that decision makers would be willing to pay for the right to use the dominant distribution. The lower bound measures the shift in the dominant distribution that produces an efficient set with both distributions as members. It measures the amount that at least one decision maker in the preference interval would pay for the information in the dominant distribution. Other preferences might be associated with a greater willingness-to-pay. The upper bound measures the shift in the dominant distribution that results in the inferior distribution being preferred to the dominant distribution. This reflects the most that any decision maker would pay for the dominant distribution, but some decision makers would be willing to pay a smaller amount.

In this option, the user will be asked to specify the dominant distribution, the alternative distributions, a decrement size, a maximum number of iterations, and a preference interval. The decrement size and the number of iterations will define the resolution of the analysis. The product of the decrement size and the number of iterations should exceed any anticipated premiums that decision makers would be willing to pay to adopt the preferred strategy. The decrement size can be set large to identify the neighborhood of the solution, and the program will automatically refine the estimate by reducing the decrement in the final iterations. A sample output of option 6 appears in Table 4. The first line identifies the preference interval and the name of the input file for the distribution data. The dominant distribution and the size of the decrement that is iteratively subtracted from the dominant distribution are also

displayed. The names of the distributions to which the dominant distribution is compared are exhibited with their associated lower and upper bounds on the value of information. For example, the value of information in strategy "cap50" for the class of decision makers represented by the preference interval is between \$4.50 and \$8.50 per acre when strategy "t0430" is used as the prior knowledge condition.

The next option is designed to succinctly display summary information about the distributions in the choice set. It will calculate the mean, the standard deviation, the element with the highest value, the element with the lowest value, and a measure of skewness. There are three options for information display: 1) printer and screen; 2) disk and screen; or 3) screen only.

The last option simply allows the analyst to perform standard DOS commands from within the program.

DISTRIBUTION FILE FORMAT

The program has been designed to directly

read standard ASCII data files that are only minimally formatted. The file must contain on the first line the number of distributions and the number of elements making up each distribution (no particular placement within the line is necessary). Then, for each of the distributions, a line is devoted to the distribution name, with the data values on the subsequent line. The data values require no particular formatting and may appear all on one line or spread over several lines, as desired.

SOURCE CODE

Most of the program has been written in Microsoft Fortran 77 and Turbo PASCAL with the graphics component coded in Microsoft Assembly language. The program is currently in use in more than 30 states and eight foreign countries. A copy of the program can be obtained from the Department of Agricultural Economics at the University of Arkansas by sending an unformatted diskette, name, and address.

REFERENCES

- Anaman, K. A., and W. G. Boggess. "A Stochastic Dominance Analysis of Alternative Marketing Strategies for Mixed Crop Farms in North Florida." *So. J. Agr. Econ.*, 18(1986):257-66.
- Bosch, D., and V. Eidman. "Valuing Information When Risk Attitudes Are Nonneutral: An Application to Irrigation Scheduling." *Amer. J. Agr. Econ.*, 69(1987):658-68.
- Byerlee, D., and J. R. Anderson. "Risk, Utility and the Value of Information in Farmer Decision Making." *Rev. Mkt. Agr. Econ.*, 50(1982):231-45.
- Cochran, M. J., and J. W. Mjelde. "Estimating the Value of Information with Stochastic Dominance: An Application from Agricultural Crop Management." Texas Agricultural Experiment Station Technical Article TA 23245. 1987.
- Cochran, M. J. "Stochastic Dominance: The State of the Art in Agricultural Economics" in *Risk Analysis for Agricultural Production Firms: Concepts, Information Requirements and Policy Issues*. Proceedings of the 1986 meetings of Regional Project S-180. Department of Agricultural Economics, Washington State University. August 1986.
- Cochran, M. J., L. J. Robison, and W. Lodwick. "Improving the Efficiency of Stochastic Dominance Techniques Using Convex Set Stochastic Dominance." *Amer. J. Agr. Econ.*, 67(1985):289-95.
- Danok, A. B., B. A. McCarl, and T. K. White. "Machinery Selection Modeling Incorporation of Weather Variability." *Amer. J. Agr. Econ.*, 62(1980):700-08.
- Greene, C., R. Kramer, G. Norton, E. Rajotte, and R. McPherson. "An Economic Analysis of Soybean Integrated Pest Management." *Amer. J. Agr. Econ.*, 67(1985):567-72.
- Hilton, R. W. "The Determinants of Information Value: Synthesizing Some General Results." *Management Science*, 27(1981):57-64.
- Holt, M. T., and J. A. Brandt. "Combining Price Forecasting with Hedging of Hogs: An Evaluation Using Alternative Measures of Risk." *J. Futures Mkts.*, 5(1985):297-309.
- King, R. P., and D. W. Lybecker. "Flexible, Risk Oriented Marketing Strategies for Pinto Bean Producers." *West. J. Agr. Econ.*, 8(1983):124-33.
- King, R. P., and G. E. Oamek. "Risk Management by Colorado Dryland Wheat Farmers and the Elimination of the Disaster Assistance Program." *Amer. J. Agr. Econ.*, 65(1983):247-55.

- King, R. P., and L. J. Robison. "Implementation of the Interval Approach to the Measurement of Decision Maker Preferences." Research Report 418, Agricultural Experiment Station, Michigan State University. November 1981.
- Kramer, R. A., and R. D. Pope. "Participation in Farm Commodity Programs: A Stochastic Dominance Analysis." *Amer. J. Agr. Econ.*, 63(1981):119-28.
- Lemieux, C., J. Richardson, and C. Nixon. "Federal Crop Insurance vs. ASCS Disaster Assistance for Texas High Plains Cotton Producers: An Application of Whole Farm Simulation." *West. J. Agr. Econ.*, 7(1982):141-53.
- Meyer, J. "Choice Among Distributions." *J. Econ. Theory*, 14(1977):326-36.
- Rister, M. E., J. Skees, and J. R. Black. "Evaluating Use of Outlook Information in Grain Sorghum Storage Decisions." *So. J. Agr. Econ.*, 16(1984):151-58.
- Tauer, L. W. "Use of Life Insurance to Fund the Farm Purchase from Heirs." *Amer. J. Agr. Econ.*, 67(1985):60-69.
- Zacharias, T. P., and A. H. Grube. "An Economic Evaluation of Weed Control Methods Used in Combination With Crop Rotation: A Stochastic Dominance Approach." *N. Cent. J. Agr. Econ.*, 6(1984):113-20.
- Zering, K. D., C. O. McCorkle, and C. V. Moore. "The Utility of Multiple Peril Crop Insurance for Irrigated, Multiple-Crop Agriculture." *W. J. Agr. Econ.*, 12(1987):50-59.