501

# Farm marketing decisions

# under uncertainty



by

Timothy A. Payne

A Thesis Submitted to the

Graduate Faculty in Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

Major: Economics

Signatures have been redacted for privacy

Iowa State University
Ames, Iowa

1980

TABLE OF CONTENTS

# LIST OF TABLES

Page

# LIST OF FIGURES

# CHAPTER I. INTRODUCTION

The marketing decision faced by farmers each fall is important in determining the farmers' incomes. At the time of harvest each farmer is faced with a decision whether to sell his grain at that time or to store it for some period of time with the possibility that he can increase his income by doing so. If he markets his grain at the "right" time, he stands to improve his financial condition, sometimes considerably. This is evidenced by an example from the 1972-73 marketing season. The Iowa average price of soybeans at harvest time in October of 1972 was \$3.06 per bushel but by the following June the price had risen to an amazing \$10.10 per bushel. In this instance, the time of marketing was extremely important. The farmers who held their soybeans until June made a gross gain due to marketing of \$7.04/bushel - considerably more than enough to offset any additional costs they may have incurred. Certainly, this is an extreme example which quite probably could not have been foreseen ahead of time but it does show that the farmer's income is not entirely dependent upon his production decision.

It was mentioned previously that the marketing decision was based on the possibility of increasing farm revenues due to higher commodity prices at some point later in the marketing season. (It is assumed here that the farmer will market his grain prior to the next year's harvest. The time between successive harvests, then, is defined to be the "marketing season".) It is imperative that we stress the word "possibility" in the above statement for it is what lies at the crux of the matter.

If future prices were known with certainty, the marketing decision would pose no problem. Each farmer could compute his costs of storage and his opportunity costs for varying lengths of time and then choose to market his commodities at the time which would maximize his net gain. Certain knowledge of future prices is not the case, however. Farmers are faced with a competitive market for their commodities in which prices fluctuate due to the varying forces of supply and demand. It has been shown (Gonzalez-Mendez [10]) that these price fluctuations take place randomly around a value of central tendency. This value of central tendency is what Gonzales-Mendez (G-M) has labelled the break-even price (bep). It is a price which is based on the cash price at harvest time and which incorporates the farmer's opportunity and storage costs. It is, in essence, a price which will leave the farmer indifferent between selling his grain at harvest time at price $p_o$ or selling it i periods later at price $p_i = bep_i$. This simplifies the farmer's decision somewhat because he can now compare his expectations of future prices to the break-even prices. If he expects future cash prices to exceed his break-even price, then he will choose to store his grain in order to increase his income. If he expects cash prices at all points in the future to fall below the break-even price he will market his grain immediately.

These criteria for selection of the proper marketing strategy may or may not be valid depending upon the farmer's risk preference. It is inevitable that forecasting further into the future will yield riskier forecasts and this greater riskiness may well affect the farmer's marketing decision. The problem with many decision making models that include risk in their specification, however, is that they often become too cumbersome to be put to practical use by many businesses. If this is the case, selection of a risk-return model may produce theoretically sound results which, due to their complexity, are never put to practical use by the farmer. This brings us to the objective of this thesis.

A plethora of models have been advanced to deal with the problem of decisions under uncertainty and several of these are discussed within the body of this paper. The decision maker is faced with several models to choose from. There is a tradeoff of sorts between the models not unlike the risk-return tradeoff. Some models, though very thorough theoretically, involve massive amounts of computation which the farmer is either unwilling or unable to undertake. In many cases, they involve the use of complex mathematical designs which can only be solved with the aid of a computer. It is quite probable that most farm operators do not have access to facilities necessary to put these models to use. On the other hand, the farmer can make use of some very simplified models which require little computation. In most cases, however, these models are lacking theoretical justification and often produce unreliable results.

The problem that will be addressed in this thesis is how to construct a model which will yield reliable results with a minimum amount of computational effort. The work of G-M will be used as a basis from which a model reflecting sound economic analysis will be constructed in the form of some simple rules of thumb.

# CHAPTER II. REVIEW OF LITERATURE

## Linear Programming Techniques

Linear programming (LP) is a mathematical model design whereby an individual can optimize a linear objective function which is subject to linear constraints. The technique is employed extensively in agriculture, business, economics and elsewhere due to its relative ease of use. The rudiments of LP are discussed widely in the literature and will not be examined here. Let it suffice to say that the main advantage of LP over other model designs (e.g., quadratic programming) is its ease of computation. This is due to the introduction of the simplex method by Dantzig [17] which is readily programmed into the computer and which allows for fast and efficient determination of the LP instrument variables. The major drawback to LP is the fact that many problems cannot be described using only linear expressions. The game theory, maximum admissible loss and MOTAD models discussed hereafter are all linear in nature and draw upon LP in one way or another.

## Game Theory Techniques

Game theory techniques have been widely used in agricultural decision making. In this context, game theory has dealt with "games against nature," i.e., the farmer's adversary is his own physical environment. The "game" placed before him is a decision as to which action, $a^*$, to undertake from among n such actions, $a_1$, $a_2$, ..., $a_n$. Uncertainty enters the decision-making process when it becomes clear that the farmer's actions will result in different outcomes depending on the existing state of nature, $s_j$. The action chosen is dependent upon the goals of the individual decision maker. Specific game theories address themselves to the differing goals which confront the decision maker.

The Wald [33] maximin criterion specifies as its optimal action that action which will lead to the maximum minimum payoff. This simply means that, given all possible states of nature, the algorithm will choose the action which has the largest minimum payoff. While this criterion will not lead to a profit maximization, per se, it will set a floor on the payoff to be received by the decision maker. It is, therefore, a technique which is quite useful to those individuals who must maintain some minimum income level in order to stay in business.

Another game theory technique is embodied in the Savage [25] regret criterion. It operates on the assumption that individuals attempt to minimize "regret". For each state of nature, the difference between the payoff for any given action and the maximum possible payoff is computed. These differences are defined as being the regret felt as a result of not choosing the optimum $a_i$ for the existing $s_j$. The criterion here is to choose that action which minimizes the maximum regret over all states of nature.

The Hurwicz [14] pessimism-optimism criterion is basically an extension of the Wald approach but allows the decision-maker to subjectively assign probabilities to the occurrence of various states of nature. The individual assigns a number, $b_i$ ($0 \le b_i \le 1$), to each action to represent his belief that the worst possible state of nature for that action will occur. This number is his "pessimism index". The individual's "optimism index" is then $(1 - b_i)$. The "$b_i$ index" for any given action, $a_i$, is then computed as

$$(2.1)\quad b_i m_i + (1-b_i)\, M_i = b_i \text{ index}$$

where

$m_i$ = minimum payoff for action $a_i$ over all states of nature,

$M_i$ = maximum payoff for action $a_i$ over all states of nature.

The optimal Hurwicz action is chosen as the action which maximizes the $b_i$ index. It is evident that if $b_i = 1$ then the solution is equivalent to that obtained using the Wald criterion. Problems arise in this technique because of the very subjective nature of assigning numeric value to the $b_i$'s.

The final game theory technique to be considered here is the Laplace [27] criterion. It is based on the assumption that the decision-maker is completely ignorant of the probabilities of different states of nature. This allows him to assign equal probabilities to every $s_j$. The expected outcome, based on these probabilities, of each action is then computed and the action with the largest expected outcome is chosen as the optimal strategy. This model has sometimes been referred to as the "naive model" since it assumes no knowledge concerning possible states of nature. If any information is available, then it is obvious that the results of this technique do not constitute the best strategy open to the decision-maker.

The game theories cited in the literature have been used quite extensively because of the relative ease in using them. The principles underlying them are straightforward and the computations involved in reaching a decision are all linear. These techniques suffer, however, from at least two major drawbacks. First, most of the game theories (as used in agriculture) are based on the assumption that nature is malevolent. That nature is consciously attempting to do its worst seems to be a rather far-fetched assumption. Second, the game theory models rely quite heavily on subjective data. Forming a Hurwicz pessimism index, for example, involves trying to quantify a great deal of subjective, as well as objective, information into a single number - a task which is difficult at best.

## Maximum Admissible Loss Approach

Boussard [ 3 ] and Boussard and Petit [ 4 ] have developed a linear model which could be considered a "safety-first" model. The idea is to maximize farm revenues subject to a minimum acceptable income. The model defines a "maximum admissible loss", L, as the difference between expected income and the minimum income needed to finance a "bare bones" consumption. This comprises the first safety constraint and can be written as

(2.2) $L = E - MINI$

where

$E$ = the income function to be maximized,

MINI = the minimum income necessary to meet basic consumption levels.

The remaining constraints involve the individual activities. A "possible unitary loss" for activity i, $FL_i$, is defined as the difference between expected income for that activity, $E(a_i)$, and the income that would be obtained if the most unfavorable of conditions prevailed. The assumption of the model is that the activity safety constraints will be satisfied if an activity's focal loss, $FL_i x_i$, does not exceed a specified fraction, 1/K, of the maximum admissible loss. (The value of K is dependent upon the distribution of activity incomes. Boussard and Petit [ 4 ] have shown that when net revenues per unit are normally distributed $K^2 = n^*$ is a reasonable approximation where $n^*$ is the number of activities in the optimal plan. This, of course, poses a problem in that $n^*$ isn't known until the system is solved although we do know that $n^* \leq$ number of constraints.)

The activity safety constraints are then set up as

$$(2.3) \quad FL_i x_i \leq L/K \qquad i = 1, 2, \ldots, n$$

where

$x_i$ = level of activity i.

If the activity incomes can be assumed to be independently normally distributed, Kennedy and Francisco [15] have shown that these constraints effectively restrict the chance that total income will fall below MINI to a maximum probability determined by the decision maker.

This approach is somewhat of a "black box" approach in that some of the underlying assumptions of the model are debatable.

> "This [model] rests on hypotheses that are more or less debatable, as has been seen. Thus, there can be no question of basing on these hypotheses the conclusions that can be drawn from it. Hence, these conclusions must rest on the forecasting value of the model - that is to say, on its ability to provide results which, under given conditions, reproduce to a reasonable approximation, the behavior of the farmers placed under these conditions" [3].

In defense of the technique, at least one of these assumptions - that the variance of activity incomes are independent of the levels of chosen activities - has been discarded with no pronounced effect on the results [15]. Some of the other assumptions still remain in question, however. For example, the assumption of independence between activity incomes is certainly open to debate. Another problem arises in selection of the possible unitary losses. How the farmer decides what his income will be if nature does its worst is difficult to determine empirically because he has no way of deciding what "nature's worst" really is.

## Minimization of Total Absolute Deviations Model

The minimization of total absolute deviations (MOTAD) model has been constructed by Hazell [13] as an alternative to the E-V quadratic models normally postulated. The specification of the model allows solutions to be obtained using standard LP algorithms. Instead of determining an efficient E-V frontier, MOTAD derives an efficient expected income - mean absolute income deviation (or E-A) frontier. The mean absolute income deviation, A, is defined as follows. (See Hazell [13].)

$$(2.4) \quad A = \frac{1}{s} \sum_{h=1}^{s} \left| \sum_{j=1}^{n} (c_{hj} - g_j)\, x_j \right|$$

where

$s$ = sample size of activity gross margins

$c_{hj}$ = the h-th observed value of the gross margin for activity j, $h = 1, 2, \ldots, s$

$g_j$ = mean activity gross margin for activity j

$x_j$ = level of activity j

$n$ = number of activities.

Derivation of the efficient E-A frontier is accomplished by minimizing A as E is varied parametrically from zero to its maximum value. The problem is then analogous to the portfolio problem in that the individual must choose that point along the E-A efficient frontier which maximizes utility.

Since this model will be dealt with in greater detail elsewhere in this paper, let it suffice to say that the major advantage of MOTAD is in its ability to be specified using LP rather than quadratic programming.

## Monte Carlo Programming

Monte Carlo programming is not a mathematical technique, per se, but rather a search process. Using the computer, a large number of activity combinations, chosen at random, can be scanned to find the optimal portfolio. The activity portfolios are first tested for feasibility. If a particular activity combination fails to satisfy the constraints of the decision problem, it is discarded. If the values assigned to activity levels do satisfy the set of constraints, then these values are plugged into the objective function. The computer can then compare all of the objective function results and choose the "optimal" value. While this optimum may or may not coincide with the true optimum value of the objective function (subject to the given constraints), the results should yield a close approximation given a large enough set of possible portfolios.

The advantages of this technique are primarily twofold [1]. The first is that virtually any specification of the objective function can be handled. This is particularly useful when the objective function is nonlinear and the necessary advanced algorithms are either unavailable or too cumbersome. In these cases, the Monte Carlo solution is probably a reliable substitute for the "true" solution. The other major advantage is that the technique allows for use of integer constraints. Since many optimization problems require the use of integer constraints, this is a particularly important advantage.

Monte Carlo programming is not a viable technique, however, when exact optima are required. Also, the technique suffers from its reliance upon the computer. An optimization problem with a large number of possible activities will undoubtedly consume much of the decision-maker's budget in computer costs.

## Bayesian Decision Models

The Bayesian approach to decisions under uncertainty involves the use of a great deal of probability information - both subjective and objective. The approach is also divided into two separate models, DATA and NONDATA. (Note here that it is possible to specify the NONDATA model using subjective data only.) Briefly, the difference between the two is that the DATA model utilizes forecasts of future states of nature or other sample information whereas the NONDATA model does not. Let us begin by examining the components of the Bayesian NONDATA model.

The NONDATA model is composed of five components. From Ladd and Williams [19] they are:

(a) a set of actions available to the decision maker,

(b) a set of states of nature that affect the outcomes of each action,

(c) a set of payoffs for each combination of action and state of nature,

(d) a "prior" probability distribution over states of nature that shows the decision maker's estimate of the probability of occurrence of each state of nature, and

(e) a rule for selecting one action from among all of the available actions.

The set of actions available to the decision maker consists of those physical acts that the decision maker can feasibly undertake in an attempt to achieve a desired outcome. The problem arises in the fact that these actions have different consequences depending on the existing state of nature. The Bayesian procedure attempts to enumerate all possible combinations of actions and states of nature and to identify the corresponding outcomes (or payoffs).

Once the payoffs associated with each action-state pair are defined, probabilities are assigned to the occurrence of each state of nature. The decision maker now has some idea of his chances of receiving a particular payoff given that he is undertaking a particular course of action. With this information at hand, the decision maker can derive expected payoffs from each action. The decision criterion, according to the Bayesian procedure, is to select the action which yields the highest expected payoff. Let us look a little more closely at the steps involved in the model.

The available actions, the $a_i$, are those actions which the decision maker feels are feasible and relevant to the problem at hand.

The states of nature, the $s_j$, confronting the individual are those conditions that he feels will have an effect on the outcome of his actions. For example, the real estate developer who has to decide how many homes to build must consider the possible states of the economy at the time his homes are ready to be put on the market. As this example suggests, states of nature can often be continuous variables. If this is the case, the distribution of states may, as one alternative, be broken into a suitable number of intervals to simplify the procedure.

The payoffs for each action-state pair, the $G_{ij}$, are now computed. It is assumed here that, given a specific action and state of nature, the corresponding payoff is known, i.e., for any given $a_i$ and $s_j$ there is only one $G_{ij}$.

The threads that weave all of the preceding components together are the prior probabilities. These probabilities, the $P(s_j)$, are assigned to each state of nature by the decision maker and, in the case of the "nondata prior", are that individual's personal, subjective estimates. They may be based on experience, research, relevant theory or simply casual observation. In many research applications prior probability vectors used are those known as "data priors". These are obtained from past frequency distributions of the states of nature.

With all of the above information at hand, the decision maker can now compute expected payoffs and make his decision. The expected payoff of action i is computed as

$$(2.5)\quad EP(a_i) = \sum_j G_{ij}\, P(s_j) \qquad (i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, m)$$

Following the Bayesian criteria leads the decision maker to choose that action which satisfies the equation

$$(2.6)\quad EP(a_i^*) = \max_i EP(a_i)$$

i.e., choose that action which maximizes the expected payoff.

The Bayesian DATA model contains all of the components of the NONDATA model plus forecast information. If we let $Z_k$ be the forecast that state of nature $s_k$ will occur, then $P(Z_k)$ is the unconditional probability that $s_k$ is forecast. At this point, conditional probability also becomes a factor in the computations. We define $P(Z_k|s_j)$ to be the probability that state of nature $s_k$ is forecast given that $s_j$ actually occurs. We now combine the conditional probabilities with the prior probabilities in order to obtain posterior probabilities. This is done using equation (2.7):

$$(2.7)\quad P(s_j|Z_k) = \frac{P(s_j)\ P(Z_k|s_j)}{P(Z_k)} \qquad j,k = 1,\ 2,\ \ldots,\ m$$

where

$$P(Z_k) = \sum_j P(s_j)\ P(Z_k|s_j)$$

The posterior probabilities, equation (2.7), give the probability that state of nature $s_j$ actually occurs given that state $s_k$ was forecast.

Again, as in the NONDATA model, expected payoffs are computed. The equation used for this purpose is shown in (2.8):

$$(2.8)\quad EP(a_{ik}) = \Sigma\ G_{ij}\ P(s_j|Z_k)$$

where

$EP(a_{ik})$ = the expected payoff of action $a_i$ given forecast $Z_k$

The Bayesian strategy is to maximize expected payoff for the given forecast, i.e.,

$$(2.9)\quad EP(a^*_{ik}) = \max_i EP(a_{ik}) \qquad i = 1,\ 2,\ \ldots,\ n;\ k = 1,\ 2,\ \ldots,\ m$$

From the decision-making methods described earlier, the disadvantages of this model are readily apparent. They are 1) the reliance upon subjective probability and 2) choosing an action with a criteria that doesn't consider the higher moments of the probability distribution. It is probably ironic that this method's greatest weakness is also its greatest strength. The ability of the Bayesian procedure to incorporate subjective data into the model quantitatively allows the decision maker to feel that he is making his own decision. Because he provides his own unique priors, it seems to him that he has some say in the decision process.

## Portfolio Analysis

The planning problem encountered under conditions of uncertainty can also be handled by using portfolio analysis. Markowitz [21] authored the definitive treatise on the subject in 1959 and a great deal of research in the area has been done in the interim. The basic assumption of the model is that the individual considers only the first two moments of a risky portfolio in making a decision concerning the optimal portfolio to hold. Although some [2,20] have criticized the appropriateness of this assumption, the technique is useful in obtaining at least some information relevant to the decision-making process. A conclusion reached by the model is that, assuming the returns on all assets in the portfolio are not perfectly positively correlated, diversification of investments will yield a lower standard deviation of return than a straight weighted average of their individual standard deviations. Indeed, it is quite possible to obtain a portfolio variance which is less than any of the individual security variances contained within it.

A simple portfolio model containing only two risky assets can be constructed as such:

return equation:

$$(2.10) \quad R_p = A_x E(R_x) + A_y E(R_y)$$

where

$R_p$ = the expected return on the portfolio

$R_x, R_y$ = the return on risky assets x and y, respectively

$E(R_x)$, $E(R_y)$ = the expected returns on assets x and y, respectively

$A_x, A_y$ = the proportions of x and y included in the portfolio

and where $A_x + A_y = 1$

risk equation:

$$(2.11) \quad \sigma_p^2 = A_x^2 \sigma_x^2 + A_y^2 \sigma_y^2 + 2A_x A_y \text{ Cov } (x, y)$$

where

$\sigma_p^2$ = the variance of the portfolio returns

$\sigma_x^2$ = the variance of returns on asset x

$\sigma_y^2$ = the variance of returns on asset y

Cov (x,y) = the covariance between the returns on x and y

and since cov $(x,y) = \rho_{xy} \sigma_x \sigma_y$

where

$\rho_{xy}$ = the correlation between the returns on x and y;

$(-1 \leq \rho \leq 1)$

therefore

$$(2.12) \quad \sigma_p^2 = A_x^2 \sigma_x^2 + A_y^2 \sigma_y^2 + 2A_x A_y \rho_{xy} \sigma_x \sigma_y$$

(At this point, we can now show that the proposition concerning the standard deviation of the portfolio when $\rho \neq 1$ stated earlier can be proven. We stated that diversification of investments, assuming $\rho \neq 1$, will yield a lower standard deviation of return than would be expected by taking a weighted average of the standard deviations where the weights used are $A_x$ and $A_y$. The weighted average standard deviation for the portfolio is

$$(2.13) \quad \sigma_{pw} = A_x \sigma_x + A_y \sigma_y$$

and the standard deviation of the diversified portfolio is

$$(2.14) \quad \sigma_p = \sqrt{A_x^2 \sigma_x^2 + A_y^2 \sigma_y^2 + 2A_x A_y \rho_{xy} \sigma_x \sigma_y}$$

If $\rho = 1$, then

$$(2.15) \quad \sigma_p = \sqrt{A_x^2 \sigma_x^2 + A_y^2 \sigma_y^2 + 2A_x A_x \sigma_x \sigma_y}$$

$$= A_x \sigma_x + A_y \sigma_y$$

which is equivalent to (2.13) above. If $\rho < 1$ then

$$(2.16) \quad \sigma_p = \sqrt{A_x^2 \sigma_x^2 + A_y^2 \sigma_y^2 + 2p\, A_x A_y \sigma_x \sigma_y}$$

which, since $2\rho\, A_x A_y \sigma_x \sigma_y < 2A_x A_y \sigma_x \sigma_y$, is less than $A_x \sigma_x + A_y \sigma_y$.)

Obtaining the appropriate first and second derivatives of equations (2.10) and (2.12) allows us to represent the relationship between mean and variance of the portfolio as seen in Figure 2.1 (assuming $E(R_y) > E(R_x)$ and $\sigma_y^2 > \sigma_x^2$.)

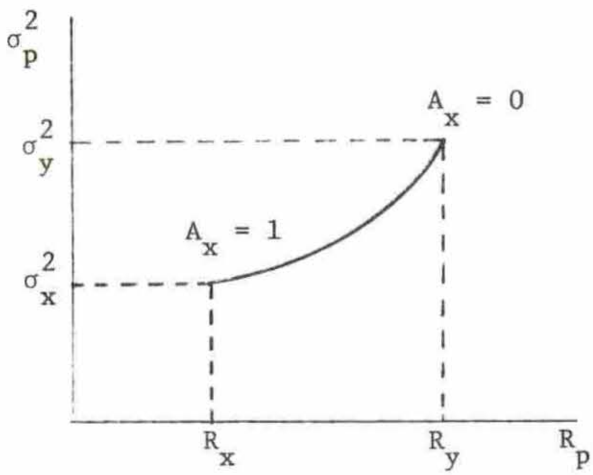


Figure 2.1. Combinations of $\sigma_p^2$ and $R_p$ afforded by varying $A_x$

The line connecting $(R_x, \sigma_x^2)$ and $(R_y, \sigma_y^2)$ represents the various combinations of $\sigma_p^2$ and $R_p$ afforded by varying the value of $A_x$. For example, when $A_x = 1$, $\sigma_p^2 = \sigma_x^2$ and $R_p = R_x$. The curvature of the line in the graph is dependent upon the value of the correlation coefficient. When $\rho = 1$, the curvature is very slight. When $\rho = -1$, the curve "bends" down to the $R_p$ axis indicating that there exists some combination of x and y which will completely diversify away all riskiness. These two cases are shown in Figures 2.2a and 2.2b in the more familiar mean-standard deviation space where the lines become straight rather than curved.

The efficient set of portfolios can be defined as those portfolios which have the highest expected return for any given level of risk.

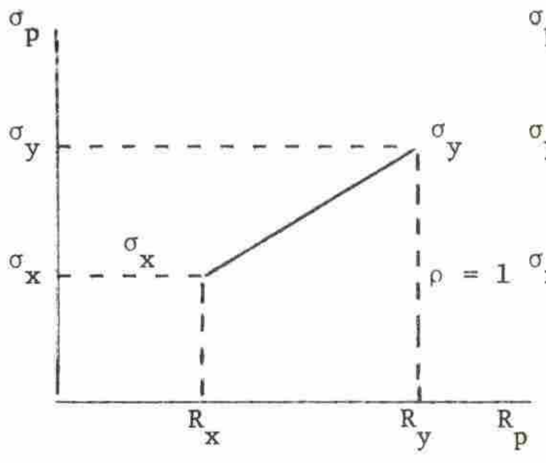


Figure 2.2a. Combinations of E & V when $\rho = 1$

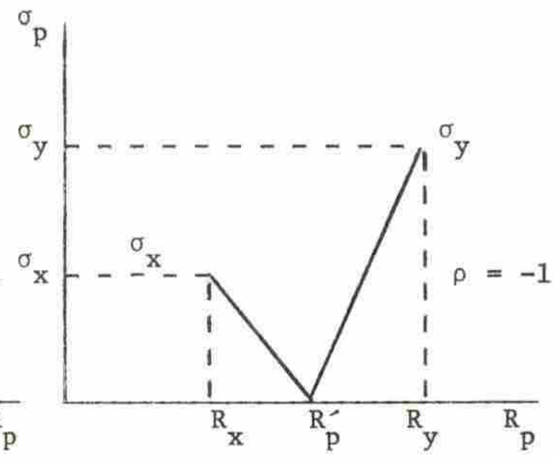


Figure 2.2b. Combinations of E & V when $\rho = -1$

For example, the line $R'_p\sigma_y$ in Figure 2.2b would represent the efficient set of portfolios given $\rho = -1$. The line $\sigma_x R'_p$ would be inefficient because the individual could obtain portfolios with higher expected returns that subject him to no more risk simply by moving in a horizontal direction to the line $R'_p\sigma_y$.

Given an efficient set (or efficient frontier) the individual must decide which point along the frontier he prefers. This is done using the familiar concept of utility. Assuming again a world with two risky assets, x and y, the individual's utility function can be expressed as $u = u(x,y)$ and, by setting $du/dx = 0$, we can obtain his family of indifference curves. If it can be assumed that the individual is risk-averse, the indifference curves can be proven to be concave downward [21,31]. This implies that the direction of increasing utility is to the southeast (see Figure 2.3). The decision maker will maximize utility by choosing that portfolio which allows him to obtain the indifference curve representing the largest level of utility.

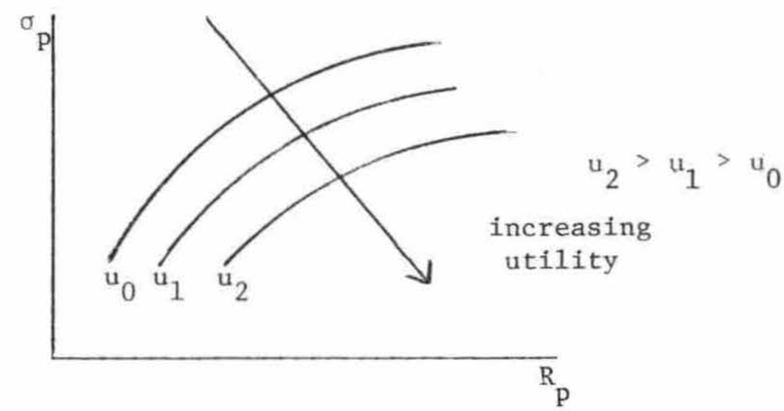


Figure 2.3. Utility map

This occurs at the point where the efficient frontier is just tangent to one of the indifference curves ($P^*$ in Figure 2.4). This result does not occur in the case of a risk-neutral or risk-preferring individual. In these cases, a corner solution will exist. It can be

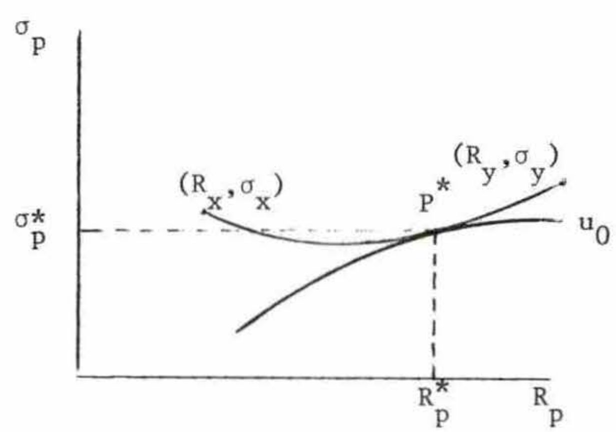


Figure 2.4. Maximization of utility

shown that risk-neutral and risk-preferring individuals have the indifference maps which are represented by Figures 2.5a and 2.5b, respectively [21].

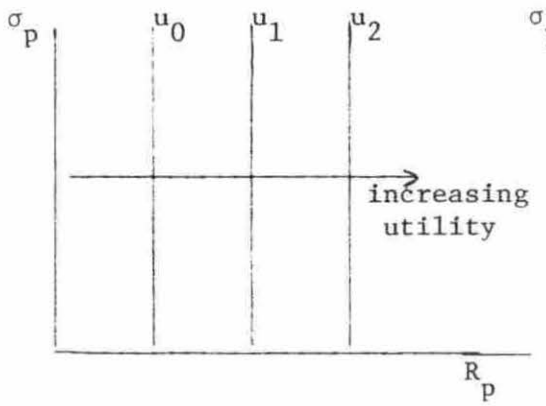


Figure 2.5a. Indifference curves under assumption of risk-neutrality

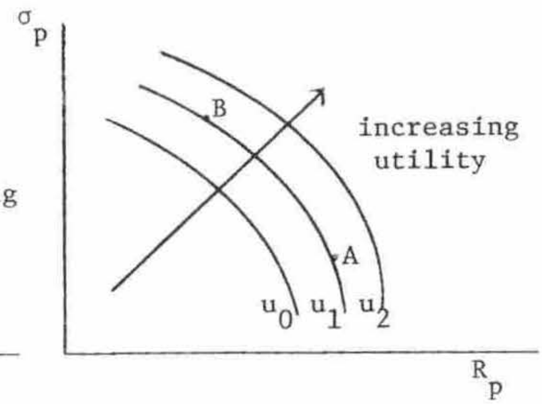


Figure 2.5b. Indifference curves under assumption of risk-preference

As can be seen by Figure 2.5a, the risk-neutral individual is concerned only with expected return and will be indifferent between two assets of differing riskiness as long as they both yield the same expected return. This individual will always choose a portfolio consisting entirely of asset y as seen in Figure 2.6a (at point $P^*$).

Figure 2.5b shows that the risk-lover will be indifferent between an asset with (relatively) high return and low risk (point A) and an asset with (relatively) low return and high risk (point B). With the direction of increasing utility being to the northeast, this also yields a portfolio consisting entirely of asset y (point $P^*$ in Figure 2.6b).

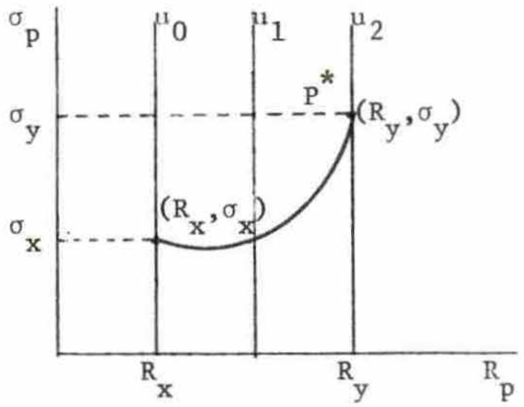


Figure 2.6a. Utility-maximizing portfolio given risk-neutrality

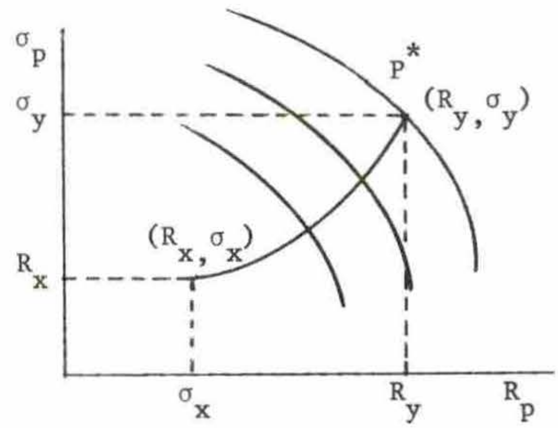


Figure 2.6b. Utility-maximizing portfolio given risk-preference

The portfolio analysis can be expanded to include non-risky assets as well as any number of risky assets. Its most obvious application is to stock portfolios but it has been applied to many varied fields; whole-farm planning [1], monetary theory [9,28,29] and insurance [24], to name a few. The portfolio problem is amenable to solution by several mathematical models. The most commonly used of these is quadratic programming. This is the natural choice given the standard specification of the utility function as quadratic. Markowitz [21] has shown, however, that dynamic programming, Monte Carlo techniques, gradient methods and even linear programming can be employed as computing techniques under certain circumstances.

The problems associated with portfolio analysis are primarily threefold. The first is that, in most situations, the portfolio selection must be done using quadratic programming. This causes problems because of the large amounts of data inherently needed to employ quadratic programming as well as deficiencies in the computing algorithm itself [1,13]. The second problem is one alluded to earlier. Much criticism has been levelled against the method for taking into account only mean and variance. While higher moments of the utility function can be worked into the analysis, they are done so at a considerable computational cost [ 1 ]. The final major argument levelled against the technique is one common to all methods which rely on utility maximization - how do you ascertain the individual's indifference curves? Although some progress has been made in this area [12], the derivation of indifference curves and the specification of utility functions remains a lengthy and somewhat dubious procedure [see Appendix]. It should be pointed out, in defense of this method, that it probably isn't necessary for the analyst to specify individual indifference curves. It is sufficient to present the decision maker with his available options and allow him to choose the action(s) which he prefers. In this way, the decision maker is maximizing his utility without verbally or otherwise communicating his utility function to the researcher.

# CHAPTER III. APPLICATION OF DECISION MODELS TO THE MARKETING DECISION

The problem to be considered in this paper deals with the farm marketing decision. In particular, we are interested in furthering the work done by Hector Eduardo Gonzalez-Mendez in his doctoral dissertation, "Grain marketing gains in Iowa and the use of price forecasting models - a Bayesian decision approach" [10]. In his paper, Gonzalez-Mendez (G-M) formulated a Bayesian model which the farmer could use to determine, at harvest time, whether he should sell his grain or store it for future sale.

## Bayesian Model

G-M assumes that the crop is available for marketing immediately after harvest. The farmer's only choice at that point is whether to sell or store. For him to make the decision to store, he must foresee some gain from postponing sale of the grain in question.

In analyzing this problem, G-M broke the marketing season up into discrete one-month intervals. These are indicated by the subscripts "h" for the month of harvest and $ph_i$ ($i = 1, 2, \ldots, 10$) for the i post-harvest months of the marketing season. For example, if November is the harvesting month for corn, then a variable with subscript "h" represents that variable's value in November. Likewise, $ph_1$ signifies December, $ph_2$ signifies January and so on. He then assumes that there is some price pattern which will leave the farmer indifferent between selling at harvest or storing until any later month. This set of prices is defined to be the set of break-even prices, the $bep_i$. These are calculated as:

$$(3.1) \quad bep_i = CP(h)(1+r)^i + SC_i \qquad i = 1, 2, \ldots, 10$$

where

$CP(h)$ = harvest time cash price

$r$ = rate of interest per ph period

$SC_i$ = costs of storage from h to $ph_i$

This says that the farmer must be at least compensated for his storage costs plus some premium for his risk-taking in order to induce him to store his crop. (In the G-M thesis, this risk premium is quantified using annual rates of interest paid by farmers. It seems probable that a more suitable measure of the risk premium could be obtained by using a rate of interest which is more closely associated with the farmer's opportunity rate of interest. This would be some interest rate which the decision maker could realize by selling his crop at harvest time and investing in an alternative financial instrument.) By appeal to Muth's rational expectation hypothesis, G-M assumed, and subsequently tested, that the expectation of the future cash price would be the break-even price. This allows us to consider a distribution of possible cash prices around the $bep_i$ as seen by

$$(3.2) \quad CP(ph_i) = bep_i + u_i$$

$$E(u_i) = 0; \; E(u_i^2) = \sigma_i^2$$

$$\sigma_i^2 > \sigma_j^2 \qquad \text{for } i > j$$

The diverse elements present in the determination of actual cash prices and the size of the sample set would suggest, through the central limit theorem, that the $u_i$ are approximately normally distributed. Although the data from the 1955-1976 marketing years indicate somewhat higher kurtosis in the distribution of the $u_i$ than would be expected of a normal distribution, G-M assumed normality in order to simplify the problem. This implies, then, that the CP ($ph_i$) are normally distributed with mean $bep_i$ and variance $\sigma_i^2$. In order to define specific states of nature of the future cash prices within each $ph_i$, these distributions of cash prices were divided into five equally likely intervals, the $s_j$. The midpoint of each interval was used in determining payoffs. For example, $CP_{ij}$ is the midpoint of the j-th state of nature in $ph_i$ ($j = 1,2,3,4,5$; $i = 1,2,\ldots,10$).

The possible actions open to the farmer in any marketing period are two - either sell or store, with selling in period $ph_i$ defined as action $a_i$. As mentioned earlier, the farmer will sell only if he foresees no gain from storage. It now becomes necessary to determine the gain from each action over all states of nature. Let

$G_{ij}$ = gain from action i if $s_j$ is the state of nature prevailing in $ph_i$ where $G_{ij} = CP_{ij} - bep_i$.

Given that $CP_{ij}$ and $bep_i$ can be computed at harvest time, the decision maker can now construct a payoff (gain) matrix which represents the $G_{ij}$ for all possible combinations of $a_i$ and $s_j$ (see Figure 3.1).

| $a_i \backslash s_j$ | $s_1$ | $s_2$ | $s_3$ | $s_4$ | $s_5$ |
|---|---|---|---|---|---|
| $a_0$ | 0 | 0 | 0 | 0 | 0 |
| $a_1$ | $G_{11}$ | $G_{12}$ | $G_{13}$ | $G_{14}$ | $G_{15}$ |
| $a_2$ | $G_{21}$ | $G_{22}$ | $G_{23}$ | $G_{24}$ | $G_{25}$ |
| $a_3$ | $G_{31}$ | $G_{32}$ | $G_{33}$ | $G_{34}$ | $G_{35}$ |
| $a_4$ | $G_{41}$ | $G_{42}$ | $G_{43}$ | $G_{44}$ | $G_{45}$ |
| $a_5$ | $G_{51}$ | $G_{52}$ | $G_{53}$ | $G_{54}$ | $G_{55}$ |
| $a_6$ | $G_{61}$ | $G_{62}$ | $G_{63}$ | $G_{64}$ | $G_{65}$ |
| $a_7$ | $G_{71}$ | $G_{72}$ | $G_{73}$ | $G_{74}$ | $G_{75}$ |
| $a_8$ | $G_{81}$ | $G_{82}$ | $G_{83}$ | $G_{84}$ | $G_{85}$ |
| $a_9$ | $G_{91}$ | $G_{92}$ | $G_{93}$ | $G_{94}$ | $G_{95}$ |
| $a_{10}$ | $G_{10,1}$ | $G_{10,2}$ | $G_{10,3}$ | $G_{10,4}$ | $G_{10,5}$ |

Figure 3.1. Payoff matrix

The Bayesian procedure now requires that prior probabilities be assigned to the possible states of nature. G-M introduces four prior probability vectors which he thinks farmers may plausibly use. One characterizes a farmer who assumes all states of nature are equally likely (prior I), one assumes a pessimistic (i.e., lower prices more likely) outlook (prior P), another assumes an optimistic outlook (prior O), and the final prior probability vector (prior N) characterizes a farmer who has a strong appeal for the value of central tendency (i.e., the probability of $CP_{ij} - bep_i = 0$ is quite high). To a limited degree, G-M also incorporates data-priors into his NONDATA model. He uses historical data for corn and soybeans over the 1955-1977 period to derive estimates of the frequency distribution of prices.

In order to include a Bayesian DATA model in the analysis, G-M introduces five possible forecasting models. These include a Trend Price Model (TPM), a Moving Average Price Model (MAPM), a Two-Variable Linear Model (TVLM), a Single-Equation Reduced Form Model (SEM) and a model which utilizes futures quotes from the Chicago Board of Trade (CBT-F) as forecasts.

The Trend Price Model assumes that the difference between the price in period $ph_i$ and period h remains the same from one year to the next. The forecasted price is specified as

$$(3.3) \quad FP(ph_i)_t = CP(h)_t + [CP(ph_i)_{t-1} - CP(h)_{t-1}]$$

where the "t" and "t-1" subscripts represent year t and year t-1, respectively.

The Moving Average Price Model postulates that this year's forecasted price for period $ph_i$ can be obtained by using a five-year moving average value of the price series. The forecast is obtained using

$$(3.4) \quad FP(ph_i)_t = \sum_{y=t-5}^{t-1} [CP(ph_i)_y]/5.$$

The Two-Variable Linear Model assumes that the expected price in period $ph_i$ is a linear function of past prices in period $ph_i$, i.e.,

$$(3.5)\quad CP(ph_i)_t = c_i + d_i t + u_{it}$$

where $c_i$ and $d_i$ are least-squares estimates using data from the five years preceding the year to be forecast.

The Single-Equation Model is a reduced-form equation which combines information from both a supply and a demand function. The least-squares equation used makes $CP(ph_i)$ a function of only two predetermined variables, $Y_i$ and $Z_i$.

$$(3.6)\quad CP(ph_i)_t = a_i + b_i Y_{it} + c_i Z_{it} + e_{it}$$

where

$Y_{it}$ = consumer disposable income at the beginning of month i in marketing season t

$Z_{it}$ = stock of corn (or soybeans) on hand at the beginning of month i in marketing season t.

This equation, too, was estimated using data for the five years immediately preceding the forecast year.

The Futures Market Model uses futures prices reported by the Chicago Board of Trade to forecast prices for the months of the marketing season corresponding to the months in which futures contracts are subject to delivery. These reported prices are then adjusted to allow for the basis present between Chicago and the point in Iowa under consideration.

The final elements necessary for computation of the optimal

Bayesian DATA strategy are the conditional and posterior probabilities. The conditional probabilities measure the accuracy of the particular forecast and are necessary in computation of the posterior probabilities. In the G-M thesis, they are computed as the percentage of times that some price $CP_{ih}$ was forecast given that $CP_{ij}$ actually occurred. These probabilities, denoted as $P_r(FP_{ih}|CP_{ij})$, were computed for the 22 years prior to November of 1976.

The posterior probabilities are obtained from straightforward computations combining the conditional and prior probabilities already computed. The posterior probabilities, denoted as $P_r(CP_{ij}|FP_{ih})$, give an idea of how often $CP_{ij}$ occurs given that $CP_{ih}$ is forecast beforehand. They are computed as

$$(3.7)\quad P_r(CP_{ij}|FP_{ih}) = [P_r(CP_{ij})\ P_r(FP_{ih}|CP_{ij})]/P_r(FP_{ih})$$

where

$P_r(CP_{ij})$ = prior probability that state of nature j will occur in period i

$$P_r(FP_{ih}) = \sum_{t=1}^{5} P_r(CP_{it})\ P_r(FP_{ih}|CP_{it})$$

Given the immense amount of data manipulation afforded by all of the above computations, G-M sets out to determine (using DATA and NON-DATA strategies) which course of action will maximize expected payoffs in the 1976-77 marketing season.

The NONDATA strategy, as mentioned earlier, is to maximize expected gain without the use of forecast information. First, expected gains from each action must be computed as

$$(3.8)\quad EG(a_i) = \sum_{j=1}^{5} G_{ij}\, P_r(CP_{ij})$$

The decision maker must now choose the $a_i$ which maximizes $EG(a_i)$, i.e., choose the action which satisfies

$$(3.9)\quad EG^*(a_h) = \max_i EG(a_i)$$

The farmer, at harvest time, now chooses whether to sell or store on the basis of whether $EG(a_i) \geq 0$ for any month other than the month of harvest. If it is, he stores. If not, he sells. G-M arrived at several $EG^*(a_h)$ due to his use of four nondata priors and five data-priors. The majority of the results indicated that the farmer should store his grain for at least one month.

The DATA strategy makes use of the posterior probabilities calculated with the help of price forecasts. As in the NONDATA case, expected gains are first computed.

$$(3.10)\quad EG(a_i|FP_{ik}) = \sum_{j=1}^{5} G_{ij} P_r\, (CP_{ij}|FP_{ik})$$

where

$FP_{ik}$ is the actual price forecast.

The optimal strategy is again to maximize expected gains, i.e.,

$$(3.11)\quad EG^*(a_h|FP_{ik}) = \max_i EG(a_i|FP_{ik})$$

G-M derived optimal DATA strategies for all combinations of nondata priors, forecasting models and two commodities, corn and soybeans.

This time, however, his harvest time strategy was consistent throughout - store the grain for at least a one-month period.

The biggest problem that one encounters with the G-M thesis is its inability to react to new information that becomes available to the farmer following the month of harvest. The paper presents a model which chooses an optimal marketing date as of the time of harvest. It does not deal with changing conditions in the market which may have an impact on when the maximum gain from storage will occur. For example, a governmental program which will affect farm prices may occur later in the marketing season which could not be foreseen at the time the optimal Bayesian strategy was formulated.

It is quite possible to reestimate the model periodically throughout the marketing season to incorporate new information. Indeed, G-M proposed just that idea in his closing remarks. Although this has been done to some limited degree (see Gonzalez-Mendez and Ladd [11]), the data requirements for several combinations of priors, forecasts and commodities would be tremendous. If a suitable data bank and computer program to generate strategies could be developed, however, it would be a valuable addition to the information now provided farmers to aid in marketing decisions.

## Game Theory Techniques

### Wald maximin criterion

As pertains to the grain marketing problem cited in the G-M thesis, use of the Wald criterion would be extremely easy and straightforward. The payoffs formulated by G-M, along with a definition of all possible states of nature for each marketing period (i.e., the $s_j$'s from the thesis) would provide adequate data to choose a marketing action under the Wald guidelines. The $a_i$ chosen (i = 1, 2, ..., 10) would be the one which yielded the largest minimum gain over all states of nature. This implies, however, that unless some $a_i$, $i > 0$, has nonnegative payoffs over all states of nature, the farmer's decision will automatically be to sell at harvest time. It seems that, given any knowledge of the probabilities of the future $s_j$, this technique will not, on average, choose a marketing strategy for the farmer which will yield the maximum expected gain.

Savage regret criterion

The Savage criterion can also be implemented using the payoff matrix from the G-M thesis. If we assume, as G-M did, that five states of nature and eleven marketing actions exist, we could compute the "regret" corresponding to each of the 55 action-state pairs. The first step would be to find the maximum gain for each state of nature. For each state we would then subtract from this state's maximum gain the gains received from every other action. The results are summarized in the Regret Payoff Matrix, Figure 3.2.

The problem is now reduced to the choice of that action which has the smallest maximum regret ($G_{max_{ij}} - G_{ij}$) over all j. By doing this the farmer is supposedly minimizing the "amount of regret" that he feels from not having chosen the optimal action given the state of nature that does prevail and determines the outcome of his action.

State of Nature

| action | $s_1$ | $s_2$ | $s_3$ | $s_4$ | $s_5$ |
|---|---|---|---|---|---|
| $a_0$ | $G_{max_{i1}} - G_{01}$ | $G_{max_{i2}} - G_{02}$ | $G_{max_{i3}} - G_{03}$ | $G_{max_{i4}} - G_{04}$ | $G_{max_{i5}} - G_{05}$ |
| $a_1$ | $G_{max_{i1}} - G_{11}$ | $G_{max_{i2}} - G_{12}$ | $G_{max_{i3}} - G_{13}$ | $G_{max_{i4}} - G_{14}$ | $G_{max_{i5}} - G_{15}$ |
| $a_2$ | $G_{max_{i1}} - G_{21}$ | $G_{max_{i2}} - G_{22}$ | $G_{max_{i3}} - G_{23}$ | $G_{max_{i4}} - G_{24}$ | $G_{max_{i5}} - G_{25}$ |
| $a_3$ | $G_{max_{i1}} - G_{31}$ | $G_{max_{i2}} - G_{32}$ | $G_{max_{i3}} - G_{33}$ | $G_{max_{i4}} - G_{34}$ | $G_{max_{i5}} - G_{35}$ |
| $a_4$ | $G_{max_{i1}} - G_{41}$ | $G_{max_{i2}} - G_{42}$ | $G_{max_{i3}} - G_{43}$ | $G_{max_{i4}} - G_{44}$ | $G_{max_{i5}} - G_{45}$ |
| $a_5$ | $G_{max_{i1}} - G_{51}$ | $G_{max_{i2}} - G_{52}$ | $G_{max_{i3}} - G_{53}$ | $G_{max_{i4}} - G_{54}$ | $G_{max_{i5}} - G_{55}$ |
| $a_6$ | $G_{max_{i1}} - G_{61}$ | $G_{max_{i2}} - G_{62}$ | $G_{max_{i3}} - G_{63}$ | $G_{max_{i4}} - G_{64}$ | $G_{max_{i5}} - G_{65}$ |
| $a_7$ | $G_{max_{i1}} - G_{71}$ | $G_{max_{i2}} - G_{72}$ | $G_{max_{i3}} - G_{73}$ | $G_{max_{i4}} - G_{74}$ | $G_{max_{i5}} - G_{75}$ |
| $a_8$ | $G_{max_{i1}} - G_{81}$ | $G_{max_{i2}} - G_{82}$ | $G_{max_{i3}} - G_{83}$ | $G_{max_{i4}} - G_{84}$ | $G_{max_{i5}} - G_{85}$ |
| $a_9$ | $G_{max_{i1}} - G_{91}$ | $G_{max_{i2}} - G_{92}$ | $G_{max_{i3}} - G_{93}$ | $G_{max_{i4}} - G_{94}$ | $G_{max_{i5}} - G_{95}$ |
| $a_{10}$ | $G_{max_{i1}} - G_{10,1}$ | $G_{max_{i2}} - G_{10,2}$ | $G_{max_{i3}} - G_{10,3}$ | $G_{max_{i4}} - G_{10,4}$ | $G_{max_{i5}} - G_{10,5}$ |

Figure 3.2. Regret payoff matrix

where $G_{max_{ij}}$ = the maximum payoff over all $a_i$, given state of nature j (j = 1,2,3,4,5).

## Hurwicz pessimism-optimism criterion

Under this criterion, the decision maker looks only at the best and the worst possible outcomes from any given action. To the worst outcome he assigns the coefficient $b_i$ ($0 \le b_i \le 1$) to indicate his belief (subjective probability) that the state of nature which gives rise to that outcome will actually occur. This is, in a sense, similar to the formulating of prior probabilities in the Bayesian model. To the best (maximum) outcome he now assigns the coefficient $(1-b_i)$, his "optimism index". The objective is to maximize the "$b_i$ index" over all i where

(3.12) $b_i \text{ index} = b_i \min_j G_{ij} + (1-b_i) \max_j G_{ij}$

In the G-M problem this implies that the decision maker would choose from among eleven (i=0,1,2,...10) $b_i$ indices that index which was the largest. The procedure may be simplified somewhat by following the approach outlined below.

1. Using a payoff matrix which lists the gains from any action-state pair, determine the maximum and minimum gains for each action over all states of nature.
2. Determine which states of nature correspond to the maximum and minimum gains found in step (1) above.
3. For each action assign a number b, $0 < b < 1$, to denote the decision maker's level of pessimism concerning the state of nature which corresponds to the minimum gain for that action.
4. Form the $b_i$ index for each $a_i$.
5. Choose the action which yields the highest value of the $b_i$ index.

The chief advantage of this method (as well as the Bayesian model) is that it allows the farmer to interject his own beliefs concerning future states of nature into the model. This gives him the satisfaction of knowing that it is "his" model and reflects his judgment. One large flaw in the model, however, is that it considers only the worst and the best outcomes and disregards those in-between. If all payoffs except one were bunched close to the minimum payoff, then the maximum payoff would carry an inordinate amount of weight, as seen in the example below.

| $a_i$ \ $s_j$ | $s_1$ | $s_2$ | $s_3$ | $s_4$ |
|---|---|---|---|---|
| $a_1$ | 10 | 15 | 20 | 50 |
| $a_2$ | 10 | 40 | 45 | 50 |

Figure 3.3. Payoff matrix

Upon inspection, it becomes clear that action $a_2$ is superior to $a_1$ because, for every $s_j$, the payoff is greater than or equal to that available under $a_1$. Because $s_1$ is the state of nature prevailing when the minimum payoff is encountered for both $a_1$ and $a_2$, it can be presumed that the decision maker will assign equal values of b to both actions. This, then, implies that the individual is indifferent between the two actions as seen by equations (3.13) and (3.14):

$$(3.13) \quad \text{action } a_1: \; b_1 \text{ index} = 10\, b_1 + 50\,(1-b_1)$$
$$= 50 - 40\, b_1$$

$$(3.14)\quad \text{action } a_2\text{:}\quad b_2 \text{ index} = 10\ b_2 + 50\ (1-b_2)$$
$$= 50 - 40\ b_2$$

where

$$b_1 = b_2$$

It is, therefore, obvious that the Hurwicz criterion is capable of producing some nonsensical results.

Laplace criterion

The Laplace criterion, which is based on the "principle of insufficient reason," is considered a game theory technique by some [6,34] and so will be included here. As we have seen earlier, it is assumed under the Laplace criterion that all states of nature are equally likely. Under this assumption, the decision problem boils down to choosing the action with the highest "average" return. All that the decision maker must do is add up the gains for action $a_i$ over all $s_j$ and divide that sum by the number of states of nature in order to get an "average gain" for action $a_i$. The decision becomes to choose that action with the largest average gain. As regards the marketing decision, the Laplace criterion is to choose the marketing strategy which will give the farmer the largest average gain to storage over all states of nature. While this criterion is extremely easy to use it neither provides nor utilizes sufficient data for the farmer to make a completely informed decision. If, for instance, the farmer has a good idea of the probable states of nature which will occur in the future, why not use this information in helping him plan his marketing strategy?

## Maximum Admissible Loss Approach

The Maximum Admissible Loss Approach (MALA), as developed by Boussard and Petit [ 4 ] is a linear programming model which is concerned with the production, rather than the marketing, decision. The constraints imposed upon the MALA model are concerned with keeping the farmer from losing more than a specified amount, given the worst conceivable state of nature for the chosen combination of activities.

The first step in specifying the MALA model is to determine what the maximum admissible loss, L, is to be. As seen before, L is calculated as the difference between expected income and the minimum permitted income, MINI. Assume first that the farmer has a fixed amount of, let's say, corn on hand at harvest time. It is obvious that he has the option of selling the corn at that time and "losing" nothing. (A "loss" here refers to the possibility, due to conditions of uncertainty, that the price of corn may fall below the break-even price during future marketing periods. If the farmer is making his decision at harvest time, then his current corn price is equal to the break-even price by definition and he stands to make no "gain" or "loss".) This is to say that if the farmer has corn on hand and can sell it at harvest time, why should he be willing to accept anything less than this amount which he will receive with certainty? Thus, the minimum acceptable gain is equal to zero, i.e., MINI = 0.

The model can now be specified as

(3.15) $\max \Sigma \ EG(a_i) \ X_i$

(3.16) s.t. $FL_i \ X_i \leq L/K$ For all i

and

(3.17) $\Sigma X_i = 1$

where

$EG(a_i)$ = the expected gain for activity i over all states of nature

$X_i$ = percentage of the corn crop sold in marketing period i

$FL_i X_i$ = focal loss for activity i

L = maximum admissible loss

K = a constant $\geq 1$

The value of L can be computed as

(3.18) $L = \Sigma EG(a_i) \ X_i - MINI$

But as we have seen, MINI = 0 and, therefore

(3.19) $L = \Sigma EG(a_i) \ X_i - 0$

$= \Sigma EG(a_i) \ X_i$

To complete the model, the values of the $FL_i$ can be obtained using

(3.20) $FL_i = EG(a_i) - G_{mi}$ for all i, i = 1, 2, ..., n

where

$G_{mi}$ = minimum gain for activity i over all states of nature.

Now, given appropriate data for cash prices, forecasted prices and the components making up the break-even prices, the model can be solved using linear programming techniques.

The problem that is encountered with this model in the marketing framework is that it leads to internal inconsistencies. This can be seen as follows:

Define (3.21) $\max_i EG(a_i) = EG(a_m)$

where $a_m$ denotes the action which yields the maximum expected gain.

Then (3.22) $\Sigma EG(a_i)\ X_i \leq EG(a_m)\ \Sigma X_i$

because $EG(a_i) \leq EG(a_m)$ for all i

and (3.23) $EG(a_m)\ \Sigma X_i = EG(a_m)$

because $\Sigma X_i = 1$ from (3.17)

Therefore, to maximize $\Sigma EG(a_i)\ X_i$ we must set

(3.24) $EG(a_i) = EG(a_m)$

and (3.25) $X_m = 1$

This yields

$$(3.26) \quad \max \Sigma EG(a_i)\ X_i = EG(a_m)\ X_m = EG(a_m)$$

From (3.26) we can see that the MALA model constrains the decision maker to marketing all of his grain in one period. This, in turn, leads to the aforementioned inconsistency in the model.

If we assume, as seems plausible, that there exists some state of nature for every marketing action i, i > 0, such that

(3.27) $CP_i < bep_i$

then (3.28) $G_i < 0$

because $CP_i - bep_i = G_i$

This implies that

(3.29) $G_{mi} < 0.$

Therefore, from (3.20)

(3.30) $FL_i > EG(a_i).$

Now, given that $FL_i > EG(a_i)$ from (3.30) and that $L = EG(a_i)$ or $L = 0$ from (3.19) (as $X_i = 1$ or $X_i = 0$, respectively), this implies

(3.31) $FL_i > L \quad \text{for } i > 0.$

Therefore

(3.32) $FL_iX_i > L/K \quad \text{when } X_i = X_m = 1$

because $K \geq 1$

and (3.33) $FL_iX_i \leq L/K \quad \text{when } X_i \neq X_m$

because $X_i = 0$

Expression (3.32) runs contrary to (3.16) for the one action, $a_m$, that is included in the solution. The model excludes the gain-maximizing action and is, therefore, internally inconsistent.

There is, however, one exception to this dichotomy. If the optimal marketing decision is to sell at harvest time, i.e., $a_m = a_o$, then $L = EG(a_o) = 0$. Since, in this case

(3.34) $EG(a_o) = G_{mo} = 0$

then (3.35) $FL_o = EG(a_o) - G_{mo} = 0$

Plugging this back into the activity safety constraints yields

(3.36) $FL_o \leq L/K$

or $0 \leq 0/K$

or $0 \leq 0.$

This is the only marketing decision which will satisfy the model. The MALA model, therefore, is not very well suited for use in determining an optimal marketing strategy.

Portfolio Analysis and MOTAD

Because portfolio analysis and MOTAD embody essentially the same idea in different mathematical specifications, we will investigate the application of both techniques in this section.

The quadratic programming model used in portfolio analysis, as we have seen, assumes that the individual makes decisions under uncertainty on the basis of the expected return and the variance of return of every action (or bundle of actions) which confronts him. This choice criteria is valid if the decision maker possesses an E-V utility function. The quadratic model usually assumes that the individual is risk averse, i.e., $\partial E/\partial V > 0$ along any indifference curve and that E must increase at an increasing rate with higher V, i.e., $\partial E^2/\partial^2 V > 0$.

Given that the above assumptions hold, the individual can then narrow down his choice among those actions for which variance of return is minimized for a given expected return. This, then, defines the set of "efficient" actions.

As applied to the marketing decision, the parametric quadratic programming model becomes

(3.37) minimize $V_i$

(3.38) such that $EG(a_i) = \lambda \quad (0 \le \lambda \le \infty)$

where

$V_i$ = variance of gain $g_i$ $(i = 1, 2, \ldots, n)$

$$= \frac{1}{s-1} \sum_{t=1}^{s} (g_{it} - \bar{g}_i)^2 = \frac{1}{s-1} \sum_{t=1}^{s} [(CP_{it} - bep_{it}) - (\overline{CP_i - bep_i})]^2$$

$EG(a_i) = FP_i - bep_i$

This model assumes, of course, that the decision maker has in place a forecasting model by which he can obtain the $FP_i$. These could be of the same form as those presented by G-M or of any other form specified by the farmer. The model also assumes that the farmer intends to market his entire stock of grain at one time.

Solution of the quadratic programming problem yields the farmer a listing of all efficient actions available to him. He now has a "menu" of actions open to him from which he can choose an optimal marketing strategy. (This assumes that the decision maker has full knowledge of his own indifference curves. If so, he would choose that efficient marketing action which would allow him to obtain the indifference curve representing the greatest attainable expected utility - point A in Figure 3.4). Since specification of individual indifference curves is beyond the scope of this paper, we would obtain the efficient frontier and allow the farmer to subjectively make his final choice.

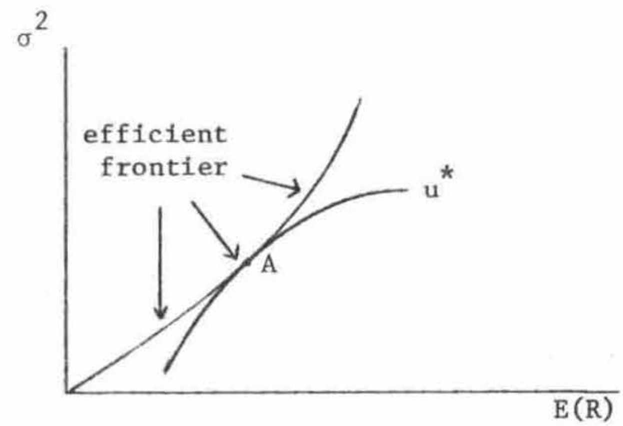


Figure 3.4. Maximization of expected utility

A problem with this choice criterion is that the marketing decision involves cross-time, rather than cross-sectional, considerations. The classic portfolio problem is concerned with choosing an optimal strategy (composed of many separate actions) at a given point in time. The problem we are faced with involves choosing the optimal time to exercise one action, namely, to sell. Since expected gains normally change over time (due to changes in forecasted prices, storage costs and desired rates of return) it is probably not prudent for the farmer to choose his marketing strategy based on the efficient frontier derived at harvest time. He can, however, use the information received from the harvest time quadratic programming solutions to choose whether he should sell or store his grain at that time. If he is willing to take the risk associated with future gains, the farmer should store his grain until a later date. If the farmer is extremely risk averse, he may choose to market his crop now even if there is a chance for significant gains later. If his decision is to store the grain, he will again be faced with the storage decision one month later. At that point, he must reformulate the problem with new price, interest rate and storage cost data and with i now varied from $i = 2, 3, \ldots, n$. He must continue this procedure every month until he chooses to sell.

Thus far in this analysis we have assumed that the farmer markets all of his grain at one time. It is quite possible (indeed probable) that the farmer allows multiple marketing actions to be considered. This may be the case for a pair of reasons. The first is that most farmers need to have at least a minimal income flow in order to meet their business and personal expenses. To the extent that other sources of income fall short of meeting this minimum level, farmers must sell grain periodically throughout the marketing season to generate adequate income. The second reason for multiple marketings is the one that portfolio analysis was originally developed to address. The farmer may want to diversify away some of the risk involved with future grain price fluctuations. If we consider marketing actions to be "assets" with given mean returns and variances of return, then we again have precisely the portfolio problem. Allowing for more than one marketing action does, however, complicate the decision problem considerably.

Given the above assumptions concerning the possible number of $a_i$ to be chosen, the optimization problem becomes:

$$(3.39)\quad \text{minimize } V = \sum_{i=1}^{n} \sum_{k=1}^{n} X_i X_k \sigma_{ik}$$

such that

$$(3.40) \quad \sum_{i=1}^{n} E(g_i)X_i = \lambda \quad (\lambda = 0 \text{ to unbounded})$$

where $g_i = CP_i - bep_i$

$$(3.41) \quad \rho(FP_i X_i) \geq y_i \qquad (\text{for all } i,\ i = 0, 1, \ldots, n)$$

$$(3.42) \quad \sum_{i=0}^{n} X_i \leq X_T$$

$$(3.43) \quad X_i \geq 0,\ Y_i \geq 0,\ 0 \leq \rho \leq 1$$

where

$X_i$ = the amount of grain sold in period i (in bu.)

$\sigma_{ik}$ = the covariance of gains between the i-th and k-th activities when $i \neq k$ and the variance of gains for the i-th activity when $i = k$

$y_i$ = the minimum acceptable income in period i

$X_T$ = total amount of grain on hand (in bu.)

n = the number of activities

$\rho < 1$ to allow for possibility that $CP_i < FP_i$.

Let us first point out that if the farmer is not faced with constraints to provide for an adequate income stream (due to other farm receipts, investment income, etc.) then equation (3.41) can readily be dropped from the model.

We now require some way to determine the variances and covariances, the $\sigma_{ik}$, of the $g_i$ in order to calculate V. Hazell [13 ] suggests that subjective parameter values for $\sigma_{ik}$ can be incorporated into the model in certain circumstances. However, the low likelihood that any farmer would be able to explicitly state numbers for these interrelationships makes this method undesirable. Using standard statistical techniques allows us to estimate V using historical data:

$$(3.44) \quad V = \sum_{i=1}^{n} \sum_{k=1}^{n} X_i X_k \left[ \frac{1}{s-1} \sum_{t=1}^{s} (g_{it} - \bar{g}_i)(g_{kt} - \bar{g}_k) \right]$$

Equation (3.44) can now replace equation (3.37) in the quadratic model.

Solution of the programming problem leads to derivation of the efficient frontier. With multiple marketings now possible, the farmer is free to choose a mixture of marketing actions, subject to the income constraints, which he feels is in his best interest. As before, however, we would expect that forecasted gains would change throughout the course of the marketing season. It becomes necessary, then, to reformulate the problem at periodic intervals throughout the year.

If a new efficient frontier is derived every month, it becomes possible to drop the income constraints altogether. We assume here that the farmer knows, at the time the efficient frontier is derived, how much money he will need in that marketing period. It is then possible for him to sell as much grain as necessary to obtain that needed level of income. With a base income for the month already secured, the farmer can now look at the marketing problem for the remainder of his grain as though it were free of income constraints.

As was stated in the review of the literature, the main drawback of portfolio analysis is that it relies on the use of quadratic programming. This causes a problem in that the quadratic programming algorithms found on most computers require large amounts of variance-covariance data and they may well be complicated. Hazell [13] has developed an approach which approximates the quadratic approach yet can be solved with linear programming techniques. It is this minimization of total absolute deviations (MOTAD) approach to which we will now turn our attention.

MOTAD

If we assume that the same sample data used in equation (3.44) are available to us now, the mean absolute gain deviation (m.a.d.) can be computed as:

$$(3.45) \quad A = \frac{1}{s} \sum_{t=1}^{s} \left| \sum_{i=1}^{n} (g_{it} - \bar{g}_i) X_i \right|$$

where

A = an unbiased estimator of the population m.a.d.

"A" is an alternative to V in measuring the dispersion of the $g_i$ parameter and has the advantage of being linear in nature rather than quadratic. It seems reasonable to assume that the decision maker can use A just as readily as he can use V in deciding between various combinations of risk and return. The problem now becomes how to derive an efficient E-A frontier from which the individual can make his decision.

Since, from equation (3.45), it is apparent that 1/s is constant, it is sufficient to minimize sA subject to the constraints (3.40) - (3.43) in order to derive an efficient E-A frontier. It becomes necessary, however, to convert sA to a form that is suitable for a linear programming model.

Define $Z_t$ as

$$(3.46)\quad Z_t = \sum_{i=1}^{n} g_{it} X_i - \sum_{i=1}^{n} \bar{g}_i X_i \qquad (\text{for all } t,\ t = 1,2,\ldots s)$$

such that (3.47) $Z_t = Z_t^+ - Z_t^-$

where $Z_t^+, Z_t^- \geq 0$

If we can now define $Z_t^+$ and $Z_t^-$ in some way such that one or the other is zero, then

$$(3.48)\quad |Z_t| = Z_t^+ + Z_t^-.$$

This can be done by defining $Z_t^+$ to be

$$(3.49)\quad Z_t^+ = \left|\sum_{i=1}^{n} (g_{it} - \bar{g}_i) X_i\right|$$

when $\sum_{i=1}^{n} (g_{it} - \bar{g}_i) X_i$ is positive and $Z_t^+$ is zero otherwise. Therefore, $\sum_{t=1}^{s} Z_t^+$ is the sum of the absolute values of the positive gain deviations about the sample mean gains. Likewise, let

$$(3.50)\quad Z_t^- = \left|\sum_{i=1}^{n} (g_{it} - \bar{g}_i) X_i\right|$$

when $\sum_{i=1}^{n} (g_{it} - \bar{g}_i) X_i$ is negative and $Z_t^-$ is zero otherwise. Thus, $\sum_{t=1}^{s} Z_t^-$ is the sum of the absolute values of the negative gain deviations about the sample mean gains.

If we set up the minimization problem as outlined earlier, it would look like

$$(3.51)\quad \text{minimize } sA = \Sigma(Z_t^+ + Z_t^-)$$

such that

$$(3.52) \quad \sum_{i=1}^{n} (g_{it} - \bar{g}_i)\, X_i - Z_t^+ + Z_t^- = 0 \qquad \text{(for all t, t=1,2,...s)}$$

and subject to constraints (3.40) - (3.43).

It is obvious, however, that

$$\sum_{t=1}^{s} Z_t^+ = \sum_{t=1}^{s} Z_t^-$$

since the sum of the positive and negative deviations about the mean must be zero. Therefore, minimization of either $Z_t^+$ or $Z_t^-$ is sufficient. This is equivalent to minimizing $\frac{1}{2}$ sA rather than sA. If we choose to minimize only the sum of the absolute values of the negative deviations, $\Sigma Z_t^-$, the model becomes

$$(3.53) \quad \text{minimize } \sum_{t=1}^{s} Z_t^-$$

such that

$$(3.54) \quad \sum_{i=1}^{n} (g_{it} - \bar{g}_i)\, X_i + Z_t^- \geq 0 \qquad \text{(for all t, t=1,2,...s)}$$

and subject to constraints (3.40) - (3.43). Again, if we assume that the efficient frontier is reformulated every month, we can drop the income constraints from consideration.

As has been stressed many times before, the chief advantage of the MOTAD approach is that it uses a linear programming algorithm rather than the quadratic programming procedure necessary for solution of the portfolio analysis formulation. Since many decision makers can gain access to a computerized LP package, whereas they may not be able to find a quadratic package, this factor becomes of prime importance.

The logical question which arises at this point is "How do the results of the two models compare?" This is a valid question since, on average, we would not expect both models to yield equivalent results.

The relevant theory behind both of these models states that the individual will select the assets in his portfolio on the basis of those assets' expected return and their riskiness of return. This implies a tradeoff between E and V or, alternatively, between E and SD (standard deviation). Since the values for expected return, E, are the same for both models, we will rule this out as a source of variation in the results and concentrate on the measure of dispersion used in each model as the source of error.

If we assume that the distribution of gains for each $a_i$ are at least approximately normally distributed, Davies and Pearson [5] have shown that the population standard deviation can be estimated using

$$SD = A\left(\frac{\pi s}{2(s-1)}\right)^{\frac{1}{2}} \text{ where } s = \text{sample size.}$$

Since $(\pi S/2(s-1))^{1/2}$ is a constant and A is the sample m.a.d., it is apparent that MOTAD generates efficient frontiers for this estimate of the population SD. As seen from equation (3.44), quadratic programming uses the conventional estimate of the population SD and, therefore, produces efficient frontiers for that measure. The differences in the reliability of the results of the two models now boils down to the reliability of each of these estimators.

Hazell [13] states that both estimators are unbiased but that there is a difference in the degree of efficiency afforded by each. Let us point out to begin with that the sample SD is the minimum variance estimator, i.e., it is the "most efficient". Hazell notes that Fisher [ 7 ] has shown that, for "large" samples, the m.a.d. is approximately 88 percent as efficient as the sample SD in estimating the population SD. Hazell goes on to show that for any sample size greater than 4 or 5, the sample m.a.d. is at least 85 percent as efficient as the sample SD. However, it should also be noted that the MOTAD model considers only variations in gains and does not look at covariations between gains. This may lead to further inefficiencies in the MOTAD model.

Aside from the oft-stated fact that MOTAD utilizes linear programming techniques and in the light of MOTAD's relative inefficiency, what are its principle advantages? Hazell [13] states that they are twofold:

> "First, it (MOTAD) may lead to much smaller problems for complex farm organizations. The quadratic programming model generally invokes m+n constraints and real activities, but the MOTAD model formulation ... requires only m+s+1 constraints and m+s real activities. Second, while quadratic programming does provide dual information on the marginal values of constraints and activities, these values do not hold over any specified intervals. The MOTAD model is therefore better adapted for post-optimality analysis."

(While we are not quite sure what Hazell means when he says "these values do not hold over any specified intervals", we must assume that he intends to state that the range over which the dual values held cannot be specified a priori. Let us also note here that

$m$ = the number of technical constraints

$n$ = the number of activities

and $s$ = the number of observations on cash prices.)

In conclusion, it is obvious that both quadratic programming and MOTAD have certain advantages over the other. Selection of the appropriate model will thus depend upon the objectives of the model-builder and the resources available to him.

# CHAPTER IV. RESULTS

The stated objective of this paper was to develop some basic rules of thumb that could easily be used by the farm operator to aid him in his grain marketing decision. We have looked at several decision-making models and have related them to the problem at hand. It was clearly evident that the portfolio model was not suited for our purposes because it involved the use of quadratic programming techniques which are not widely available and which require excessive amounts of data. The Monte Carlo and MOTAD models will also be dismissed because, while they don't require quadratic programming, they do call for a large number of computations to be performed, presumably using a computer. The game theory models have not been considered for the empirical work because, for the most part, they lack the theoretical justification that Bayesian procedures possess (Halter & Dean, 12, pp. 90-93). Finally, the maximum admissible loss approach was earlier proven to be internally inconsistent when used in the marketing framework. The empirical work will consist of simplifying the procedure used by G-M to obtain the posterior probability distributions and also to eliminate the need for breaking the distribution of states of nature into discrete units.

Our procedure involves deriving the means and variances of the posterior distributions under the assumption that the true states of nature and the forecasted states of nature are continuously distributed. If we further assume that historical cash prices (the states of nature) and the forecasted prices are approximately jointly distributed as a bivariate normal, it can be shown that the posterior distribution is also normally distributed. The rudiments of this derivation follow hereafter. (For a more detailed derivation, see Morrison, 23, pp. 84-97.)

Given that states of nature, s, and forecasted states of nature, z, are jointly normally distributed, their joint density function is given by the expression

$$(4.1)\quad f(s,z) = \frac{1}{2\pi\sigma_s\sigma_z\sqrt{1-\rho^2}}\ \exp\left[-\frac{1}{2(1-\rho^2)}\left(u^2 - \rho uv + v^2\right)\right]$$

where

$$u = \frac{z-\mu_z}{\sigma_z} \qquad \text{and } \rho = \text{correlation coefficient between s and z}$$

$$v = \frac{s-\mu_s}{\sigma_s}$$

In order to arrive at an equation for the posterior density function we must note that

$$(4.2)\quad f(s|z) = \frac{f(s,z)}{f(z)}$$

where $f(z) = \frac{1}{\sigma_z\sqrt{2\pi}} \exp\left[-\frac{u^2}{2}\right]$ if we assume that z is a normally distributed random variable

At this point, we should discuss the plausibility of assuming z to be normally distributed. G-M showed in his dissertation that all of the forecasting models other than the MAPM model had values of central tendency not significantly different from the break-even prices for all months of the marketing season and for both commodities. He also showed that the variances of the forecasts were not significantly different than the variances of the corresponding cash prices for all but the MAPM model (G-M,10, pp. 121-25). In order to determine whether the forecasted prices are normally distributed, therefore, we must measure the skewness and kurtosis of each distribution.

Skewness is a measure of the degree of deviation from symmetry for a given distribution. Since we are interested with the distribution of forecasted prices about the $bep_i$, we will be concerned with measuring the skewness of the $u_{ij}$ distribution where $u_{ij} = z_{ij} - bep_i$. ($z_{ij}$ refers to the forecasted price for the i-th month using the j-th forecasting model). The usual test of skewness is obtained by calculating the Pearsonian coefficient of skewness. Snedecor and Cochran [26] have shown, however, that it is not strictly applicable to distributions whose numbers of observations are less than 150. Mood, Graybill and Boes [22] have derived a measure of skewness which can be applied to small samples. Their measure of skewness, s, is defined as s = (mean-median)/(standard deviation). s takes on a value of zero when the distribution is completely symmetric bell-shaped (such as a normal distribution) and $-1 \leq s \leq 1$. The values of s for all $z_{ij}$ and for both commodities were calculated and appear in Tables 4.1 - 4.4.

Kolstoe [16] describes an absolute value of s greater than 0.20 as indicating a "moderate" amount of skewness. With the exception of the MAPM model (and perhaps the CBT-F model) virtually all of the

Table 4.1. Test of skewness and kurtosis of the $u_{ij}$ ($u_{ij} = FP_j(ph_i) - bep_i$) distributions of the months of the marketing season for the TPM model

| Month of marketing | Skewness (s) | | Kurtosis (a) | |
|---|---|---|---|---|
| | Corn | Soybeans | Corn | Soybeans |
| November | - | -0.1933 | - | 0.6533 |
| December | -0.1540 | -0.1096 | 0.8058 | 0.5175 |
| January | -0.1344 | -0.1633 | 0.6736 | 0.4967 |
| February | -0.1440 | -0.0505 | 0.6002 | 0.4591 |
| March | -0.1933 | -0.0302 | 0.5650 | 0.4354 |
| April | -0.2206 | -0.0366 | 0.5236 | 0.5165 |
| May | -0.1973 | 0.0373 | 0.5299 | 0.4611 |
| June | -0.0471 | 0.0969 | 0.5056 | 0.4512 |
| July | 0.0426 | 0.0561 | 0.5622 | 0.5055 |
| August | 0.1527 | 0.1664 | 0.5879 | 0.5049 |
| September | 0.0259 | - | 0.6266 | - |

| | | |
|---|---|---|
| Null Hypothesis | $H_o$: $s = 0$ | $H_o$: $a = 0.80792$ |
| Tolerance | $-1 \leq s \leq 1$ | Upper 1% = 0.9001 |
| | | Lower 1% = 0.6950 |

Table 4.2. Test of skewness and kurtosis of the $u_{ij}$ ($u_{ij} = FP_j(ph_i) - bep_i$) distributions of the months of the marketing season for the MAPM model

| Month of marketing | Skewness (s) | | Kurtosis (a) | |
|---|---|---|---|---|
| | Corn | Soybeans | Corn | Soybeans |
| November | - | -0.2039 | - | 0.5865 |
| December | -0.3036 | -0.1879 | 0.6072 | 0.5891 |
| January | -0.2534 | -0.2231 | 0.6073 | 0.6068 |
| February | -0.2810 | -0.2157 | 0.6106 | 0.6001 |
| March | -0.2968 | -0.2264 | 0.6201 | 0.6057 |
| April | -0.3082 | -0.2862 | 0.6330 | 0.6156 |
| May | -0.3043 | -0.2164 | 0.6289 | 0.5969 |
| June | -0.2809 | -0.1531 | 0.6255 | 0.5787 |
| July | -0.2618 | -0.1811 | 0.6264 | 0.5715 |
| August | -0.2560 | -0.2448 | 0.6270 | 0.5856 |
| September | -0.2960 | - | 0.6251 | - |

| | | |
|---|---|---|
| Null Hypothesis | $H_o$: s = 0 | $H_o$: a = 0.80792 |
| Tolerance | $-1 \leq s \leq 1$ | Upper 1% = 0.9001 |
| | | Lower 1% = 0.6950 |

Table 4.3. Test of skewness and kurtosis of the $u_{ij}$ ($u_{ij} = FP_j(ph_i) - bep_i$) distributions of the months of the marketing season for the TVLM model

| Month of marketing | Skewness (s) | | Kurtosis (a) | |
|---|---|---|---|---|
| | Corn | Soybeans | Corn | Soybeans |
| November | - | -0.0359 | - | 0.5576 |
| December | 0.0881 | 0.0049 | 0.5930 | 0.5642 |
| January | 0.0140 | -0.0644 | 0.6045 | 0.5729 |
| February | 0.0095 | -0.0280 | 0.6034 | 0.5847 |
| March | 0.0315 | -0.0343 | 0.6201 | 0.6135 |
| April | -0.0139 | -0.0748 | 0.5589 | 0.6178 |
| May | 0.0027 | 0.0976 | 0.5553 | 0.7004 |
| June | 0.0923 | 0.1812 | 0.5843 | 0.5825 |
| July | 0.1183 | 0.0932 | 0.5929 | 0.6539 |
| August | 0.3028 | - | 0.5739 | - |

| | | |
|---|---|---|
| Null Hypothesis | $H_o$: $s = 0$ | $H_o$: $a = 0.80792$ |
| Tolerance | $-1 \leq s \leq 1$ | Upper 1% = 0.9001 |
| | | Lower 1% = 0.6950 |

Table 4.4. Test of skewness and kurtosis of the $u_{ij}$ ($u_{ij} = FP_j(ph_i) - bep_i$) distributions of the months of the marketing season for the SEM and CBT-F models

| Month of marketing | SEM Model Skewness (s) Corn | Kurtosis (a) Corn |
|---|---|---|
| November | - | - |
| December | 0.0796 | 0.5943 |
| January | 0.0240 | 0.6006 |
| February | 0.0309 | 0.5972 |
| March | 0.0056 | 0.5870 |
| April | -0.0186 | 0.5615 |
| May | 0.0062 | 0.5608 |
| June | 0.0723 | 0.5858 |
| July | 0.2164 | 0.5927 |
| August | 0.3043 | 0.5660 |
| September | 0.2517 | 0.6370 |

CBT-F Model

| Month of marketing | Skewness (s) | | Kurtosis (a) | |
|---|---|---|---|---|
| | Corn | Soybeans | Corn | Soybeans |
| November | - | 0.0779 | - | 0.5133 |
| December | 0.2328 | - | 0.6745 | - |
| January | - | -0.0681 | - | 0.6182 |
| February | - | - | - | - |
| March | 0.1056 | -0.2264 | 0.7073 | 0.6759 |
| April | - | - | - | - |
| May | 0.1043 | -0.3538 | 0.6990 | 0.7121 |
| June | - | - | - | - |
| July | -0.4495 | -0.3899 | 0.6889 | 0.7669 |

| | Skewness | Kurtosis |
|---|---|---|
| Null Hypothesis | $H_o$: $s = 0$ | $H_o$: $a = 0.80792$ |
| Tolerance | $-1 \leq s \leq 1$ | Upper 1% = 0.9001<br>Lower 1% = 0.6950 |

calculated S values fall below this level. Because of the successful testing of the $u_{ij}$ means and variances, we fail to recognize significant amounts of skewness in the forecasted prices except in the case of the MAPM model.

Also reported in Tables 4.1 - 4.4 is a value for a, the measure of kurtosis. Kurtosis is a measure of the relative flatness or peakedness of a distribution. R. C. Geary [ 8 ] has developed a criterion for testing kurtosis in samples containing less than 200 observations. He computes a as a = (mean absolute deviation)/(standard deviation). For 21 observations, the expected value of a is 0.80792. Distributions that are more peaked (leptokurtic) than the normal distribution are characterized by lower values of a while flatter (platykurtic) distributions have larger values. These a coefficients were computed for all the $u_{ij}$ distributions for each commodity.

In virtually all cases, the value of a is significantly below its expected value. This implies that the distributions of forecasted prices about the $bep_i$ are more peaked than would be true of a normal distribution. G-M has shown [13, p. 97] that the $u_i$ distribution for cash prices is also leptokurtic and he surmises that this may be due to the inclusion of some highly abnormal marketing years in the sample. The corresponding abnormal prices tend to fall into the tails of the $u_i$ distribution which gives these observations disproportionate weight. It would seem that the effect of these abnormal prices would be transmitted to the distribution of the forecasted prices since they are used directly in the computation of most of the forecasts.

While it seems that the forecasted price distributions are more peaked than would be expected of a normal distribution, we feel that the computational benefits derived from the assumption of normality may well outweigh the problems. Therefore, we can continue the derivation as

$$(4.3)\quad f(s|z) = \frac{1}{2\pi\sigma_s\sigma_z\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}(u^2-2\rho uv+v^2)\right]$$

$$\div \frac{1}{\sigma_z\sqrt{2\pi}} \exp\left[-\frac{u^2}{2}\right]$$

$$= \frac{1}{\sigma_s\sqrt{2\pi}(\sqrt{1-\rho^2})} \exp\left[-\frac{1}{2(1-\rho^2)}(v^2 - 2\rho uv + \rho^2u^2)\right]$$

$$= \frac{1}{\sigma_s\sqrt{2\pi}(\sqrt{1-\rho^2})} \exp\left[-\frac{1}{2}\left(\frac{v-\rho u}{\sqrt{1-\rho^2}}\right)^2\right]$$

$$= \frac{1}{\sigma_s\sqrt{2\pi}(\sqrt{1-\rho^2})} \exp\left[-\frac{1}{2}\left(\frac{\frac{s-\mu_s}{\sigma_s} - \rho\frac{z-\mu_z}{\sigma_z}}{\sqrt{1-\rho^2}}\right)^2\right]$$

$$= \frac{1}{\sigma_s\sqrt{2\pi}(\sqrt{1-\rho^2})} \exp\left[-\frac{1}{2}\left(\frac{s-\mu_s-\rho(\sigma_s/\sigma_z)(z-\mu_z)}{\sigma_s\sqrt{1-\rho^2}}\right)^2\right]$$

This is the probability density function for a normal random variable with mean

$$(4.4) \quad \mu_{s|z} = \mu_s + \rho(\sigma_s/\sigma_z)(z - \mu_z)$$

and variance

$$(4.5) \quad \sigma^2_{s|z} = \sigma^2_s \ (1 - \rho^2)$$

If we take $\bar{s}$ to approximate $\mu_s$ and $\bar{z}$ to approximate $\mu_z$ and we also allow various marketing periods to be identified by subscript i, then we can compute the posterior means and variances as

$$(4.6) \quad \mu_{s_i|z_i} = \bar{s}_i + \rho_i \ (\sigma_{s_i}/\sigma_{z_i})(z_i - \bar{z}_i)$$

$$(4.7) \quad \sigma^2_{s_i|z_i} = \sigma^2_{s_i} \ (1 - \rho_i)^2$$

Having this information for the expected value of the posterior distribution, we can now circumvent the tedious procedure of computing the value of $f(s|z)$ for several values of s and then finding the expected value. (We can, therefore, omit the computations required by equations (3.7) to (3.11).) It also allows us to obtain the mean value of the continuous distribution without having to break the distribution of states of nature into discrete units.

Before computing the parameters of $f(s|z)$ we can simplify expression (4.6) somewhat by noting that

$$(4.8) \quad \rho_i \ (\sigma_{s_i}/\sigma_{z_i}) = \beta_i$$

where $(s_i - \bar{s}) = \beta_i(z_i - \bar{z}_i) + \varepsilon_i$

$\beta_i$ is a simple regression coefficient. Therefore, computation of

$\mu_{s_i|z_i}$ can be obtained using

(4.9) $\mu_{s_i|z_i} = \bar{s}_i + \beta_i(z_i - \bar{z}_i)$,

If we take historical cash prices to define the prior distribution of states of nature and historical forecast data to define the distribution of forecasts, we can now make straightforward computations of the posterior means and variances.

The first step in the procedure involves the computation of the $\beta_{ijk}$ for all combinations of forecasts, months of the marketing season and commodities. For this purpose we have used the data provided by G-M in his earlier doctoral work (G-M,10, pp. 64-65, 67-68, 72-73, 78, 85, 170-173).

Simple linear regressions of the form

$$(CP_{ik} - \overline{CP}_{ik}) = \beta_{ijk}\ (z_{ijk} - \bar{z}_{ijk}) + \varepsilon_{ijk}$$

where

$CP_{ik}$ = cash price of commodity k in marketing period $ph_i$

$\overline{CP}_{ik}$ = mean of $CP_{ik}$ over the sample period (1955-56 to 1976-77)

$z_{ijk}$ = the forecast of $CP_{ik}$ using forecasting model j

$\bar{z}_{ijk}$ = mean $z_{ijk}$ over the sample period

were run to obtain the $\beta_{ijk}$. The G-M thesis considered two commodities, corn and soybeans, and ten marketing periods (months) for each commodity. It also made use of five forecasting models (TPM, MAPM, TVLM, SEM and CBT-F) for corn and four forecasting models (TPM, MAPM, TVLM and CBT-F) for soybeans. (The SEM model was not formulated for soybeans because of a lack of data.) Since we used all of his data in our empirical work, this yielded 77 $\beta_{ijk}$ values which are listed in Tables 4.5 - 4.9. All were different from zero at the 95 percent level of significance.

Sample statistics for $\bar{s}_i$, $\bar{z}_i$ and $\hat{\sigma}^2_{s_i}$ were obtained from the historical data. The statistic $R^2$, which was used as the measure of $\rho^2$, was taken from the regression output. In order to make direct comparisons with G-M's results, the computation of the posterior means and variances were carried out for the 1976-77 marketing season. Given values for $\bar{s}_{ik}$, $\bar{z}_{ijk}$, $\hat{\beta}_{ijk}$, $R^2_{ijk}$ and $\hat{\sigma}^2_{s_{ik}}$ as well as the actual 1976-77 forecast value, $z_{ijk}$ (see Table 4.10), $\mu_{s_{ijk}|z_{ijk}}$ and $\sigma^2_{s_{ijk}|z_{ijk}}$ can be calculated over all i, j and k. These values for the means and variances can be found in Tables 4.5 - 4.9.

As we have pointed out earlier, however, the relevant statistics involve not the posterior means and variances of $f(s_i|z_i)$ but, rather, the means and variances of $f(G_i|z_i)$. In other words, the farmer is interested in his net return and not his gross return. A high cash price later in the marketing season will induce him to store his crop only if the gain in price per bushel is at least enough to offset his added storage and opportunity costs. It is apparent, therefore, that we must determine expected gains as well.

We have seen before that $G_i = CP_i - bep_i$. Since the $bep_i$ are constant for any given i in the marketing season in question, the subtraction of $bep_i$ from $CP_i$ has the effect of merely shifting the mean of $f(s_i|z_i)$ while the variance remains the same, i.e.,

$$(4.10) \quad \mu_{G_i|z_i} = (\bar{s}_i - bep_i) + \beta_i \, (z_i - \bar{z}_i)$$

Table 4.5. Summary statistics for the TPM forecasting model - 1955-56 to 1976-77

| Corn<br>Month of marketing | $\bar{S}_i$ | $\bar{Z}_i$ | $\hat{\beta}_i$ | $\hat{\rho}_i^2$ | $\hat{\sigma}_{S_i}^2$ | $E(S_i\|Z_i)$ | $E(S_i\|Z_i)^2$ | $E(G_i\|Z_i)$ | $E(G_i\|Z_i)^2$ |
|---|---|---|---|---|---|---|---|---|---|
| December | 1.3136 | 1.3059 | 0.9532 | 0.9762 | 0.3920 | 1.9848 | 0.0093 | -0.1081 | 0.0093 |
| January | 1.3273 | 1.3154 | 0.8887 | 0.9130 | 0.3827 | 2.0068 | 0.0333 | -0.1340 | 0.0333 |
| February | 1.3250 | 1.3132 | 0.8133 | 0.8228 | 0.3786 | 1.9974 | 0.0671 | -0.1876 | 0.0671 |
| March | 1.3095 | 1.3054 | 0.7681 | 0.7637 | 0.3405 | 1.9582 | 0.0805 | -0.2658 | 0.0805 |
| April | 1.3377 | 1.3264 | 0.8064 | 0.8231 | 0.2862 | 1.9857 | 0.0506 | -0.2774 | 0.0506 |
| May | 1.3841 | 1.3782 | 0.7569 | 0.7826 | 0.2959 | 2.0818 | 0.0643 | -0.2205 | 0.0643 |
| June | 1.4241 | 1.4241 | 0.6804 | 0.7089 | 0.3140 | 2.0881 | 0.0914 | -0.2535 | 0.0914 |
| July | 1.4327 | 1.4441 | 0.6465 | 0.6698 | 0.3687 | 2.1089 | 0.1217 | -0.2721 | 0.1217 |
| August | 1.4559 | 1.4764 | 0.6326 | 0.6626 | 0.5109 | 1.9832 | 0.1724 | -0.4372 | 0.1724 |
| September | 1.4095 | 1.4273 | 0.6124 | 0.6351 | 0.4231 | 1.9501 | 0.1544 | -0.5099 | 0.1544 |
| Soybeans<br>Month of marketing | | | | | | | | | |
| November | 3.0091 | 2.9986 | 0.9629 | 0.9675 | 2.1417 | 5.3214 | 0.0696 | -0.5948 | 0.0696 |
| December | 3.0705 | 3.0432 | 0.8928 | 0.9026 | 2.2461 | 4.9247 | 0.2188 | -1.0730 | 0.2188 |
| January | 3.1064 | 3.0668 | 0.7780 | 0.7995 | 2.0153 | 4.8672 | 0.4041 | -1.2073 | 0.4041 |
| February | 3.1809 | 3.1345 | 0.6470 | 0.6361 | 2.1726 | 4.6143 | 0.7906 | -1.5293 | 0.7906 |
| March | 3.2414 | 3.1541 | 0.5946 | 0.5046 | 2.4555 | 4.5114 | 1.2165 | -1.7016 | 1.2165 |
| April | 3.3409 | 3.1755 | 0.6450 | 0.4176 | 3.1397 | 4.7564 | 1.8286 | -1.5263 | 1.8286 |
| May | 3.4350 | 3.2714 | 0.4553 | 0.2565 | 3.7911 | 4.5863 | 2.8187 | -1.7663 | 2.8187 |
| June | 3.5277 | 3.4000 | 0.3653 | 0.1877 | 4.6464 | 4.8757 | 3.7743 | -1.5472 | 3.7743 |
| July | 3.3841 | 3.3254 | 0.4998 | 0.4239 | 2.8156 | 5.5255 | 1.6221 | -0.9679 | 1.6221 |
| August | 3.4909 | 3.4928 | 0.4800 | 0.4043 | 4.0641 | 5.1744 | 2.4210 | -1.3898 | 2.4210 |

Table 4.6. Summary statistics for the MAPM forecasting model - 1955-56 to 1976-77

| Corn<br>Month of marketing | $\bar{S}_i$ | $\bar{Z}_i$ | $\hat{\beta}_i$ | $\hat{\rho}_i^2$ | $\hat{\sigma}_{S_i}^2$ | $E(S_i \mid Z_i)$ | $E(S_i \mid Z_i)^2$ | $E(G_i \mid Z_i)$ | $E(G_i \mid Z_i)^2$ |
|---|---|---|---|---|---|---|---|---|---|
| December | 1.3136 | 1.1769 | 1.2787 | 0.3801 | 0.3920 | 2.4300 | 0.2430 | 0.3372 | 0.2430 |
| January | 1.3273 | 1.1896 | 1.3523 | 0.4190 | 0.3827 | 2.4854 | 0.2224 | 0.3446 | 0.2224 |
| February | 1.3250 | 1.1858 | 1.3769 | 0.4334 | 0.3786 | 2.5067 | 0.2145 | 0.3217 | 0.2145 |
| March | 1.3095 | 1.1854 | 1.4265 | 0.4588 | 0.3405 | 2.4687 | 0.1843 | 0.2447 | 0.1843 |
| April | 1.3377 | 1.2134 | 1.4352 | 0.4584 | 0.2862 | 2.3977 | 0.1550 | 0.1346 | 0.1550 |
| May | 1.3841 | 1.2573 | 1.3910 | 0.4606 | 0.2959 | 2.4923 | 0.1596 | 0.1900 | 0.1596 |
| June | 1.4241 | 1.2944 | 1.2990 | 0.4604 | 0.3140 | 2.5719 | 0.1694 | 0.2303 | 0.1694 |
| July | 1.4327 | 1.3025 | 1.1334 | 0.3627 | 0.3687 | 2.5519 | 0.2350 | 0.1709 | 0.2350 |
| August | 1.4559 | 1.3209 | 0.8763 | 0.2428 | 0.5109 | 2.5102 | 0.3869 | 0.0898 | 0.3869 |
| September | 1.4095 | 1.2860 | 0.9274 | 0.2478 | 0.4231 | 2.4018 | 0.3183 | -0.0582 | 0.3183 |
| Soybeans<br>Month of marketing | | | | | | | | | |
| November | 3.0091 | 2.6036 | 1.5918 | 0.5387 | 2.1417 | 6.2729 | 0.9880 | 0.3567 | 0.9880 |
| December | 3.0705 | 2.6611 | 1.5787 | 0.5605 | 2.2461 | 6.4030 | 0.9872 | 0.4053 | 0.9872 |
| January | 3.1064 | 2.6931 | 1.6626 | 0.6344 | 2.0153 | 6.4364 | 0.7368 | 0.3619 | 0.7368 |
| February | 3.1809 | 2.7637 | 1.5275 | 0.6159 | 2.1726 | 6.4655 | 0.8345 | 0.3219 | 0.8345 |
| March | 3.2414 | 2.8077 | 1.5929 | 0.6042 | 2.4555 | 6.6316 | 0.9719 | 0.4186 | 0.9719 |
| April | 3.3409 | 2.8417 | 1.9578 | 0.6464 | 3.1397 | 7.4059 | 1.1102 | 1.1232 | 1.1102 |
| May | 3.4350 | 2.9214 | 1.6091 | 0.5204 | 3.7911 | 7.2849 | 1.8182 | 0.9323 | 1.8182 |
| June | 3.5277 | 2.9855 | 1.3526 | 0.4404 | 4.6464 | 7.4888 | 2.6001 | 1.0659 | 2.6001 |
| July | 3.3841 | 2.8856 | 1.5029 | 0.6896 | 2.8156 | 7.4636 | 0.8740 | 0.9702 | 0.8740 |
| August | 3.4909 | 3.0027 | 1.0172 | 0.3847 | 4.0641 | 6.9649 | 2.5006 | 0.4007 | 2.5006 |

Table 4.7. Summary statistics for the TVLM forecasting model - 1955-56 to 1976-77

| Corn<br>Month of marketing | $\bar{S}_i$ | $\bar{Z}_i$ | $\hat{\beta}_i$ | $\hat{\rho}_i^2$ | $\hat{\sigma}_{S_i}^2$ | $E(S_i \vert Z_i)$ | $E(S_i \vert Z_i)^2$ | $E(G_i \vert Z_i)$ | $E(G_i \vert Z_i)^2$ |
|---|---|---|---|---|---|---|---|---|---|
| December | 1.3136 | 1.3082 | 0.6437 | 0.5740 | 0.3920 | 2.6408 | 0.1670 | 0.5480 | 0.1670 |
| January | 1.3273 | 1.3136 | 0.6681 | 0.6088 | 0.3827 | 2.6945 | 0.1497 | 0.5537 | 0.1497 |
| February | 1.3250 | 1.3118 | 0.6695 | 0.6118 | 0.3786 | 2.6896 | 0.1470 | 0.5046 | 0.1470 |
| March | 1.3095 | 1.3145 | 0.6954 | 0.6409 | 0.3405 | 2.6485 | 0.1223 | 0.4245 | 0.1223 |
| April | 1.3377 | 1.3136 | 0.7197 | 0.6880 | 0.2862 | 2.6666 | 0.0893 | 0.4035 | 0.0893 |
| May | 1.3841 | 1.3686 | 0.6987 | 0.6939 | 0.2959 | 2.7336 | 0.0906 | 0.4313 | 0.0906 |
| June | 1.4241 | 1.4241 | 0.6796 | 0.6943 | 0.3140 | 2.7329 | 0.0960 | 0.3913 | 0.0960 |
| July | 1.4327 | 1.4536 | 0.6031 | 0.5838 | 0.3687 | 2.6790 | 0.1535 | 0.2979 | 0.1535 |
| August | 1.4559 | 1.5136 | 0.5280 | 0.4834 | 0.5109 | 2.5153 | 0.2639 | 0.0949 | 0.2639 |
| Soybeans<br>Month of marketing | | | | | | | | | |
| November | 3.0091 | 2.9618 | 0.7250 | 0.6225 | 2.1417 | 5.8280 | 0.8085 | -0.0882 | 0.8085 |
| December | 3.0705 | 3.0223 | 0.7379 | 0.6464 | 2.2461 | 5.6146 | 0.7942 | -0.3831 | 0.7942 |
| January | 3.1064 | 3.0409 | 0.7929 | 0.7200 | 2.0153 | 5.6430 | 0.5643 | -0.4315 | 0.5643 |
| February | 3.1809 | 3.1296 | 0.7559 | 0.6771 | 2.1726 | 5.2599 | 0.7015 | -0.8837 | 0.7015 |
| March | 3.2414 | 3.1499 | 0.7571 | 0.5883 | 2.4555 | 4.9904 | 1.0109 | -1.2226 | 1.0109 |
| April | 3.3409 | 3.1627 | 0.8453 | 0.5190 | 3.1397 | 5.2575 | 1.5102 | -1.0253 | 1.5102 |
| May | 3.4350 | 3.2877 | 0.6452 | 0.3500 | 3.7911 | 4.7011 | 2.4642 | -1.6515 | 2.4642 |
| June | 3.5277 | 3.4341 | 0.5781 | 0.3282 | 4.6464 | 5.0573 | 3.1215 | -1.3656 | 3.1215 |
| July | 3.3841 | 3.3227 | 0.7866 | 0.7428 | 2.8156 | 6.4340 | 0.7242 | -0.0594 | 0.7242 |

Table 4.8. Summary statistics for the SEM forecasting model - 1955-56 to 1976-77

| Corn Month of marketing | $\bar{S}_i$ | $\bar{Z}_i$ | $\hat{\beta}_i$ | $\hat{\rho}_i^2$ | $\hat{\sigma}^2_{S_i}$ | $E(S_i \mid Z_i)$ | $E(S_i \mid Z_i)^2$ | $E(G_i \mid Z_i)$ | $E(G_i \mid Z_i)^2$ |
|---|---|---|---|---|---|---|---|---|---|
| December | 1.3136 | 1.3134 | 0.6361 | 0.5633 | 0.3920 | 2.6364 | 0.1712 | 0.5436 | 0.1712 |
| January | 1.3273 | 1.3179 | 0.6671 | 0.6114 | 0.3827 | 2.6833 | 0.1487 | 0.5424 | 0.1487 |
| February | 1.3250 | 1.3134 | 0.6749 | 0.6147 | 0.3786 | 2.6764 | 0.1459 | 0.4914 | 0.1459 |
| March | 1.3095 | 1.3034 | 0.7061 | 0.6505 | 0.3405 | 2.6503 | 0.1190 | 0.4263 | 0.1190 |
| April | 1.3377 | 1.3128 | 0.7158 | 0.6819 | 0.2862 | 2.6629 | 0.0910 | 0.3998 | 0.0910 |
| May | 1.3841 | 1.3703 | 0.6964 | 0.6946 | 0.2959 | 2.7311 | 0.0904 | 0.4288 | 0.0904 |
| June | 1.4241 | 1.4274 | 0.6742 | 0.6988 | 0.3140 | 2.7334 | 0.0946 | 0.3918 | 0.0946 |
| July | 1.4327 | 1.4570 | 0.6000 | 0.5913 | 0.3687 | 2.6744 | 0.1507 | 0.2934 | 0.1507 |
| August | 1.4559 | 1.5201 | 0.5271 | 0.4940 | 0.5109 | 2.5193 | 0.2585 | 0.0989 | 0.2585 |
| September | 1.4095 | 1.4977 | 0.5217 | 0.4720 | 0.4231 | 2.4318 | 0.2234 | -0.0282 | 0.2234 |

Table 4.9. Summary statistics for the CBT-F forecasting model - 1955-56 to 1976-77

| Corn<br>Month of marketing | $\bar{S}_i$ | $\bar{Z}_i$ | $\hat{\beta}_i$ | $\hat{\rho}_i^2$ | $\hat{\sigma}^2_{S_i}$ | $E(S_i \mid Z_i)$ | $E(S_i \mid Z_i)^2$ | $E(G_i \mid Z_i)$ | $E(G_i \mid Z_i)^2$ |
|---|---|---|---|---|---|---|---|---|---|
| December | 1.3136 | 1.3155 | 0.8854 | 0.9796 | 0.3920 | 2.2030 | 0.0080 | 0.1102 | 0.0080 |
| March | 1.3095 | 1.3318 | 0.7963 | 0.8899 | 0.3405 | 1.8894 | 0.0375 | -0.3346 | 0.0375 |
| May | 1.3841 | 1.4250 | 0.7443 | 0.9292 | 0.2959 | 2.0651 | 0.0210 | -0.2372 | 0.0210 |
| July | 1.4327 | 1.4168 | 0.8681 | 0.8613 | 0.3687 | 1.9911 | 0.0511 | -0.3899 | 0.0511 |
| Soybeans<br>Month of marketing | | | | | | | | | |
| November | 3.0091 | 3.1446 | 0.7111 | 0.6049 | 2.1417 | 4.9471 | 0.8462 | -0.9691 | 0.8462 |
| January | 3.1064 | 3.0655 | 0.8757 | 0.8957 | 2.0153 | 5.5360 | 0.2102 | -0.5385 | 0.2102 |
| March | 3.2414 | 3.2100 | 0.7525 | 0.6026 | 2.4555 | 5.0926 | 0.9758 | -1.1205 | 0.9758 |
| May | 3.4350 | 3.0064 | 0.8417 | 0.4471 | 3.7911 | 5.7359 | 2.0961 | -0.6167 | 2.0961 |
| July | 3.3841 | 3.0023 | 0.8487 | 0.6127 | 2.8156 | 5.6482 | 1.0905 | -0.8452 | 1.0905 |

Table 4.10. Cash prices and forecasted values for the 1976-77 marketing season

| Corn<br>Month of marketing | $CP_i$ | $bep_i$ | Forecasts<br>TPM | <br>MAPM | <br>TVLM | <br>SEM | <br>CBT-F |
|---|---|---|---|---|---|---|---|
| December | 2.22 | 2.0928 | 2.01 | 2.050 | 3.37 | 3.3930 | 2.32 |
| January | 2.31 | 2.1408 | 2.08 | 2.046 | 3.36 | 3.3505 | - |
| February | 2.30 | 2.1850 | 2.14 | 2.044 | 3.35 | 3.3158 | - |
| March | 2.28 | 2.2240 | 2.15 | 1.998 | 3.24 | 3.2023 | 2.06 |
| April | 2.27 | 2.2631 | 2.13 | 1.952 | 3.16 | 3.1641 | - |
| May | 2.20 | 2.3023 | 2.30 | 2.054 | 3.30 | 3.3045 | 2.34 |
| June | 2.09 | 2.3416 | 2.40 | 2.178 | 3.35 | 3.3694 | - |
| July | 1.84 | 2.3810 | 2.49 | 2.290 | 3.52 | 3.5265 | 2.06 |
| August | 1.56 | 2.4204 | 2.31 | 2.524 | 3.52 | 3.5375 | - |
| September | 1.59 | 2.4600 | 2.31 | 2.356 | - | 3.4572 | - |
| **Soybeans<br>Month of marketing** | | | | | | | |
| November | 6.06 | 5.9162 | 5.40 | 4.654 | 6.85 | - | 5.87 |
| December | 6.54 | 5.9977 | 5.12 | 4.772 | 6.47 | - | - |
| January | 6.71 | 6.0745 | 5.33 | 4.696 | 6.24 | - | 5.84 |
| February | 6.89 | 6.1436 | 5.35 | 4.914 | 5.88 | - | - |
| March | 7.71 | 6.2130 | 5.29 | 4.936 | 5.46 | - | 5.67 |
| April | 9.31 | 6.2827 | 5.37 | 4.918 | 5.43 | - | - |
| May | 9.20 | 6.3526 | 5.80 | 5.314 | 5.25 | - | 5.74 |
| June | 8.40 | 6.4229 | 7.09 | 5.914 | 6.08 | - | - |
| July | 6.76 | 6.4934 | 7.61 | 5.600 | 7.20 | - | 5.67 |
| August | 5.42 | 6.5642 | 7.00 | 6.418 | | - | - |

$$(4.11) \quad \sigma^2_{G_i|z_i} = \hat{\sigma}^2_{s_i} \; (1 - \rho^2_i)$$

The results of these computations for each forecasting model are also presented in Tables 4.5 - 4.9.

Comparison of Results to Those Obtained by
G-M and to Actual Gains

Comparison of our results with those of G-M is necessary in order to give some perspective to our results. There are several sources which may have an effect on the compatibility of the results. The first source of discrepancy lies in the fact that our results have been based on a Bayesian DATA approach using data priors. In his thesis, G-M postulated three models - NONDATA with nondata priors, NONDATA with data priors and DATA with nondata priors. In his DATA model, he utilized four nondata priors. I,O,N, and P, as outlined earlier. We will limit our comparisons to G-M's DATA model with nondata prior "I" because this is the prior which assumes normality of the states of nature. The second cause of differences in the results comes about because of our specification of the distribution of states of nature as continuous rather than discrete, as G-M did. Recall that G-M made the assumption that states of nature were normally distributed and that he then broke the distribution into five equally likely intervals. He proceeded to use the cash price at the midpoint of each interval when arriving at values for the expected cash price and the expected gain. This, in essence, forced him to describe the distribution of states of nature as being made up of five discrete values. Since our derivation treats the distribution as being continuous within some feasible range, this may cause a distortion in our results. Finally, we are confronted with one more possible source of discrepancy in our results. We have, as a simplifying assumption, assumed that cash prices and forecasted prices are jointly distributed as a bivariate normal and that forecasted prices are normally distributed. We have done this in order to arrive at our stated objective - to simplify the computations needed to obtain the Bayesian results. Although our statistical justification for this assumption may be lacking, we are hopeful that the ease of computation and the reliability of the results will justify this assumption.

The expected gains derived from our procedure, G-M's procedure and the actual gains for the 1976-77 marketing season are presented in Tables 4.11 and 4.12. The results for corn in Table 4.11 show that the largest actual gain observed in 1976-77 occurred in January. At that time, the farmer could have made a marketing gain of $0.169 per bushel. It should also be noted that gains show a steady decline in all months following January. In our formulation of the posterior means, the Bayesian decision (the month with the largest expected gain) was to market the corn in December in two cases and in January in two other cases. These are the two months in which actual gains were highest. Only the TPM model indicated marketing at some other time, namely at harvest. Had the farmer heeded advice from the TPM, he would have received no gain from marketing. The marketing decisions indicated by our procedure coincide with those of G-M in two instances. Both procedures yielded a Bayesian decision to market

Table 4.11. Actual gains and forecasted gains for the 1976-77 marketing season for corn

| | | Forecasts | | | |
|---|---|---|---|---|---|
| | | TPM | | MAPM | |
| Month of marketing | Actual gain[a] | G-M's model[b] | Our model | G-M's model[c] | Our model |
| December | 0.127 | -0.1658 | -0.1081 | -0.0152 | 0.3372 |
| January | 0.169 | -0.0276 | -0.1340 | -0.2034 | 0.3446 |
| February | 0.115 | 0.7767 | -0.1876 | -0.1871 | 0.3217 |
| March | 0.056 | 0.4315 | -0.2658 | -0.0284 | 0.2447 |
| April | 0.007 | -0.5560 | -0.2774 | -0.4658 | 0.1346 |
| May | -0.102 | 0.5366 | -0.2205 | -1.3551 | 0.1900 |
| June | -0.252 | -0.1277 | -0.2535 | 0.1951 | 0.2303 |
| July | -0.541 | 0.3554 | -0.2721 | 0.0000 | 0.1709 |
| August | -0.860 | -0.5024 | -0.4372 | 0.8929 | 0.0898 |
| September | -0.870 | -1.5471 | -0.5099 | -1.5471 | -0.0582 |
| Bayesian Decision | Jan. | Feb. | Nov. | Aug. | Jan. |

[a]Actual gain computed as $CP_i$-$bep_i$ using actual cash prices for 1976-77 (G-M, 10, p. 146). Values for August and September were calculated by myself.

[b](G-M, 10, p. 137).

[c](G-M, 10, p. 139).

[d](G-M, 10, p. 141).

[e](G-M, 10, p. 143).

[f](G-M, 10, p. 144).

| Forecasts | | | | | |
|---|---|---|---|---|---|
| TVLM | | SEM | | CBT-F | |
| G-M's model[d] | Our model | G-M's model[e] | Our model | G-M's model[f] | Our model |
| 0.5518 | 0.5480 | 0.4634 | 0.5436 | 0.2518 | 0.1102 |
| -0.8541 | 0.5537 | -0.8542 | 0.5424 | | |
| -0.3020 | 0.5046 | -0.3020 | 0.4914 | | |
| -0.0817 | 0.4245 | -0.2093 | 0.4263 | -0.2372 | -0.3346 |
| 0.0186 | 0.4035 | -0.0000 | 0.3998 | | |
| 0.1282 | 0.4313 | 0.1282 | 0.4288 | 0.1733 | -0.2372 |
| 0.0168 | 0.3913 | 0.0152 | 0.3918 | | |
| 0.1330 | 0.2979 | 0.1328 | 0.2934 | -0.2812 | -0.3899 |
| -0.2539 | -0.0949 | 0.1745 | 0.0989 | | |
| | | 0.0000 | -0.0282 | | |
| Dec. | Jan. | Dec. | Dec. | Dec. | Dec. |

Table 4.12. Actual gains and forecasted gains for the 1976-77 marketing season for soybeans

| | | Forecasts | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | TPM | | MAPM | | TVLM | | CBT-F | |
| Month of marketing | Actual gain[a] | G-M's model[b] | Our model | G-M's model[c] | Our model | G-M's model[d] | Our model | G-M's model[e] | Our model |
| November | 0.144 | 0.3889 | -0.5948 | -0.8409 | 0.3567 | 0.7416 | -0.0882 | 0.2803 | -0.9691 |
| December | 0.542 | 0.6456 | -1.0780 | 1.1452 | 0.4053 | -1.5256 | -0.3831 | | |
| January | 0.636 | 1.5083 | -1.2073 | -0.5143 | 0.3619 | 0.0817 | -0.4315 | 0.5381 | -0.5385 |
| February | 0.746 | 1.4141 | -1.5293 | -0.6113 | 0.3219 | -1.6322 | -0.8837 | | |
| March | 1.497 | 0.4367 | -1.7016 | -0.6760 | 0.4186 | 0.5391 | -1.2226 | 1.2119 | -1.1205 |
| April | 3.027 | 0.1951 | -1.5263 | 1.1436 | 1.1232 | 0.4584 | -1.0253 | | |
| May | 2.847 | 0.3345 | -1.7663 | -0.5143 | 0.9323 | -1.1790 | -1.6515 | -0.5372 | -0.6167 |
| June | 1.977 | 0.0000 | -1.5472 | 2.6849 | 1.0659 | -1.2505 | -1.3656 | | |
| July | 0.267 | 0.7636 | -0.9679 | 1.6392 | 0.9702 | 0.7201 | -0.0594 | -0.0152 | -0.8452 |
| August | -1.144 | 1.7722 | -1.3898 | 0.3023 | 0.4007 | | | | |
| Bayesian Decision | Apr. | Aug. | Oct. | June | Apr. | Nov. | Oct. | Mar. | Oct. |

[a]Actual gains computed as $CP_i - bep_i$ using actual cash prices for 1976-77 (G-M 10, p. 146). Value for August was computed by myself.

[b](G-M,10, p. 138).

[c](G-M,10, p. 140).

[d](G-M,10, p. 142).

[e](G-M,10, p. 144).

in December for the SEM and the CBT-F forecasting models. The differences in results for the optimal time of marketing indicated by the TPM and TVLM models were quite small. The differences in results under the MAPM model were quite large, however. G-M's procedure yielded a Bayesian solution of August for the MAPM model while ours indicated January as the optimal month. The farmer following the MAPM decision under G-M's approach would have lost \$0.86 per bushel given the break-even prices computed by G-M. Had our procedure been used, however, the farmer would have stood to make \$0.169 per bushel - a difference of \$1.029 per bushel due to the differing results of the two procedures. If a farmer would have marketed equal amounts of corn in each of the months indicated by the five models under G-M's procedure, he would have lost \$.0728 per bushel. Under our procedure, he would have made a marketing gain of \$.1184 per bushel.

As can be seen in Table 4.12, the maximum actual gain for soybeans occurred in April when marketing gains reached \$3.027 per bushel. Our procedure indicated October, the month of harvest, as the time to sell according to three of the four forecasting models. The MAPM Bayesian decision was to market soybeans in April. In contrast, G-M's procedure yielded a different month for each model. Three of the models - MAPM, which indicated June, TVLM, which indicated November and CBT-F, which indicated March - would have allowed the farmer to realize positive marketing gains. However, his TPM result, indicating August as the optimal time to sell, would have brought about a \$1.144 loss per bushel due to marketing. Equal amounts of soybeans marketed in each of the months indicated by the four models under G-M's procedure would have given the farmer an average gain of \$0.6185 per bushel. Under our procedure, the average gain would have been \$0.7568 per bushel of soybeans.

### Sensitivity of $bep_i$ to Changes in Storage Costs and Interest Rates

We stated earlier that the break-even prices were formulated using G-M's data for interest rates and storage and handling costs. It seems appropriate, in this time of volatile interest rates and double-digit inflation, that we should look at the effect that changing interest rates, cash prices and storage costs would have on the $bep_i$.

First, recall that the formulation of the breakeven prices was done using the equation

(4.12) $bep_i = CP_h\ (1+r)^i + SC_i$

The term $CP_h\ (1+r)^i$ is a measure of the farmer's opportunity cost of holding his grain rather than selling it at the time of harvest. In his dissertation, G-M used interest rates paid by farmers [30] as his measure of r. For the 1976-77 marketing season, the annual rate of interest used by G-M was 8.5 percent - or .68215 percent per month. In Tables 4.13 - 4.15, we present a listing of the changes in the $bep_i$ caused by interest rates ranging from 5.5 percent to 15 percent. We also look at how changes in $CP_o$ affect the $bep_i$. It is clear that the longer the grain is stored and the higher the $CP_o$, the larger is the effect on the $bep_i$. For instance, a 3 percent increase in the

Table 4.13. Change in $bep_i$ due to changes in the interest rate - 1976-77 marketing season (all values in \$/bu.)

| Corn: $CP_o$ = \$2.01 r<br>Month of marketing | 5.5% | 6.5% | 7.5% | 9.5% | 10.5% | 11.5% | 12.5% | 13.5% | 15% |
|---|---|---|---|---|---|---|---|---|---|
| December | -.0047 | -.0031 | -.0016 | +.0015 | +.0031 | +.0046 | +.0061 | +.0076 | +.0099 |
| January | -.0095 | -.0063 | -.0031 | +.0031 | +.0062 | +.0093 | +.0123 | +.0153 | +.0198 |
| February | -.0143 | -.0095 | -.0048 | +.0047 | +.0094 | +.0140 | +.0187 | +.0233 | +.0301 |
| March | -.0192 | -.0128 | -.0064 | +.0063 | +.0126 | +.0189 | +.0251 | +.0313 | +.0405 |
| April | -.0242 | -.0161 | -.0080 | +.0080 | +.0159 | +.0238 | +.0316 | +.0346 | +.0510 |
| May | -.0292 | -.0194 | -.0097 | +.0096 | +.0192 | +.0287 | +.0382 | +.0477 | +.0618 |
| June | -.0342 | -.0228 | -.0114 | +.0113 | +.0226 | +.0338 | +.0450 | +.0561 | +.0727 |
| July | -.0393 | -.0262 | -.0131 | +.0130 | +.0260 | +.0389 | +.0518 | +.0647 | +.0839 |
| August | -.0445 | -.0296 | -.0148 | +.0147 | +.0295 | +.0441 | +.0589 | +.0735 | +.0953 |
| September | -.0497 | -.0331 | -.0165 | +.0165 | +.0329 | +.0494 | +.0659 | +.0823 | +.1069 |
| **Soybeans: $CP_o$=\$5.80**<br>Month of marketing | | | | | | | | | |
| November | -.0136 | -.0090 | -.0045 | +.0045 | +.0089 | +.0133 | +.0176 | +.0220 | +.0284 |
| December | -.0274 | -.0182 | -.0091 | +.0090 | +.0179 | +.0268 | +.0356 | +.0443 | +.0573 |
| January | -.0413 | -.0275 | -.0137 | +.0136 | +.0271 | +.0405 | +.0538 | +.0670 | +.0867 |
| February | -.0554 | -.0368 | -.0184 | +.0183 | +.0364 | +.0544 | +.0723 | +.0901 | +.1167 |
| March | -.0697 | -.0463 | -.0231 | +.0230 | +.0458 | +.0686 | +.0912 | +.0999 | +.1472 |
| April | -.0841 | -.0559 | -.0279 | +.0278 | +.0554 | +.0830 | +.1103 | +.1376 | +.1783 |
| May | -.0987 | -.0657 | -.0328 | +.0326 | +.0652 | +.0976 | +.1298 | +.1620 | +.2100 |
| June | -.1134 | -.0755 | -.0377 | +.0376 | +.0750 | +.1124 | +.1496 | +.1867 | +.2422 |
| July | -.1283 | -.0854 | -.0427 | +.0426 | +.0851 | +.1274 | +.1697 | +.2119 | +.2750 |
| August | -.1434 | -.0955 | -.0477 | +.0476 | +.0952 | +.1427 | +.1902 | +.2375 | +.3084 |

Table 4.14. Change in $bep_i$ due to changes in the interest rate - 1976-77 marketing season (all values in \$/bu.)

| Corn: $CP_o$ = \$1.50 r<br>Month of marketing | 5.5% | 6.5% | 7.5% | 9.5% | 10.5% | 11.5% | 12.5% | 13.5% | 15% |
|---|---|---|---|---|---|---|---|---|---|
| December | -.0036 | -.0024 | -.0012 | +.0011 | +.0023 | +.0034 | +.0045 | +.0056 | +.0073 |
| January | -.0071 | -.0047 | -.0023 | +.0024 | +.0047 | +.0070 | +.0092 | +.0115 | +.0149 |
| February | -.0107 | -.0071 | -.0035 | +.0035 | +.0070 | +.0105 | +.0139 | +.0174 | +.0225 |
| March | -.0143 | -.0095 | -.0047 | +.0048 | +.0095 | +.0141 | +.0188 | +.0234 | +.0302 |
| April | -.0181 | -.0131 | -.0060 | +.0059 | +.0118 | +.0177 | +.0236 | +.0258 | +.0381 |
| May | -.0218 | -.0145 | -.0073 | +.0071 | +.0143 | +.0214 | +.0285 | +.0356 | +.0461 |
| June | -.0255 | -.0170 | -.0085 | +.0085 | +.0169 | +.0253 | +.0336 | +.0419 | +.0543 |
| July | -.0293 | -.0195 | -.0097 | +.0098 | +.0195 | +.0291 | +.0387 | +.0483 | +.0627 |
| August | -.0331 | -.0221 | -.0110 | +.0111 | +.0221 | +.0330 | +.0440 | +.0549 | +.0712 |
| September | -.0370 | -.0247 | -.0123 | +.0124 | +.0247 | +.0369 | +.0492 | +.0615 | +.0798 |
| **Soybeans: $CP_o$=\$4.80**<br>Month of marketing | | | | | | | | | |
| November | -.0113 | -.0075 | -.0038 | +.0037 | +.0074 | +.0110 | +.0146 | +.0182 | +.0235 |
| December | -.0227 | -.0151 | -.0075 | +.0075 | +.0148 | +.0222 | +.0294 | +.0367 | +.0474 |
| January | -.0343 | -.0228 | -.0114 | +.0112 | +.0224 | +.0335 | +.0445 | +.0554 | +.0717 |
| February | -.0459 | -.0305 | -.0152 | +.0151 | +.0301 | +.0451 | +.0599 | +.0746 | +.0966 |
| March | -.0577 | -.0418 | -.0191 | +.0190 | +.0380 | +.0568 | +.0755 | +.0827 | +.1219 |
| April | -.0696 | -.0462 | -.0231 | +.0230 | +.0459 | +.0687 | +.0914 | +.1139 | +.1476 |
| May | -.0817 | -.0543 | -.0271 | +.0270 | +.0539 | +.0807 | +.1074 | +.1341 | +.1738 |
| June | -.0939 | -.0625 | -.0312 | +.0311 | +.0621 | +.0930 | +.1238 | +.1545 | +.2004 |
| July | -.1062 | -.0707 | -.0353 | +.0353 | +.0704 | +.1055 | +.1405 | +.1754 | +.2276 |
| August | -.1187 | -.0791 | -.0395 | +.0394 | +.0787 | +.1181 | +.1573 | +.1965 | +.2552 |

Table 4.15. Change in $bep_i$ due to changes in the interest rate - 1976-77 marketing season (all values in \$/bu.)

| Corn: $CP_o$ = \$2.50 r<br>Month of marketing | 5.5% | 6.5% | 7.5% | 9.5% | 10.5% | 11.5% | 12.5% | 13.5% | 15% |
|---|---|---|---|---|---|---|---|---|---|
| December | -.0059 | -.0039 | -.0020 | +.0019 | +.0038 | +.0057 | +.0076 | +.0094 | +.0122 |
| January | -.0118 | -.0078 | -.0038 | +.0039 | +.0078 | +.0116 | +.0154 | +.0192 | +.0248 |
| February | -.0178 | -.0118 | -.0059 | +.0059 | +.0117 | +.0175 | +.0232 | +.0289 | +.0374 |
| March | -.0239 | -.0159 | -.0079 | +.0079 | +.0157 | +.0235 | +.0312 | +.0389 | +.0503 |
| April | -.0300 | -.0217 | -.0099 | +.0100 | +.0198 | +.0296 | +.0394 | +.0431 | +.0635 |
| May | -.0363 | -.0241 | -.0120 | +.0120 | +.0239 | +.0358 | +.0476 | +.0593 | +.0769 |
| June | -.0425 | -.0282 | -.0141 | +.0141 | +.0281 | +.0421 | +.0560 | +.0699 | +.0906 |
| July | -.0488 | -.0325 | -.0162 | +.0162 | +.0324 | +.0485 | +.0645 | +.0805 | +.1044 |
| August | -.0553 | -.0368 | -.0183 | +.0184 | +.0367 | +.0550 | +.0732 | +.0914 | +.1186 |
| September | -.0619 | -.0412 | -.0206 | +.0205 | +.0410 | +.0615 | +.0819 | +.1024 | +.1330 |
| Soybeans: $CP_o$=\$6.80<br>Month of marketing | | | | | | | | | |
| November | -.0160 | -.0106 | -.0053 | +.0053 | +.0105 | +.0156 | +.0207 | +.0258 | +.0333 |
| December | -.0322 | -.0214 | -.0106 | +.0105 | +.0210 | +.0314 | +.0417 | +.0519 | +.0672 |
| January | -.0485 | -.0322 | -.0161 | +.0159 | +.0318 | +.0475 | +.0630 | +.0785 | +.1017 |
| February | -.0650 | -.0433 | -.0216 | +.0213 | +.0426 | +.0638 | +.0848 | +.1056 | +.1368 |
| March | -.0817 | -.0592 | -.0271 | +.0269 | +.0537 | +.0804 | +.1069 | +.1171 | +.1726 |
| April | -.0986 | -.0656 | -.0327 | +.0326 | +.0650 | +.0973 | +.1294 | +.1613 | +.2091 |
| May | -.1158 | -.0770 | -.0384 | +.0383 | +.0763 | +.1144 | +.1522 | +.1900 | +.2461 |
| June | -.1330 | -.0885 | -.0442 | +.0440 | +.0879 | +.1317 | +.1754 | +.2189 | +.2839 |
| July | -.1505 | -.1002 | -.0500 | +.0499 | +.0997 | +.1494 | +.1990 | +.2484 | +.3224 |
| August | -.1681 | -.1120 | -.0560 | +.0558 | +.1116 | +.1673 | +.2229 | +.2784 | +.3615 |

interest rate (to 11.5%) would raise the bep for corn by only \$.0046 in December but would increase it by \$.0494 by September given the actual $CP_o$ of \$2.01 for corn. If the harvest time cash price had instead been \$2.50 for corn, $bep_i$ would have increased \$.0057 in December and \$.0615 in September. It is interesting to note that none of the variations in the interest rate would have affected the optimal time to market nor would they have affected the ordering of marketing decisions from best to worst. The only change would have been the nominal effect of increasing or decreasing the dollar amounts of the gains realized. The maximum amounts of these changes, given the actual $CP_o$ values of \$2.01 for corn and \$5.80 for soybeans, would have been approximately 5 cents per bushel of corn in September and approximately 14 cents per bushel of soybeans in August.

The $SC_i$ term in equation (4.12) is a measure of the farmer's actual out-of-pocket costs of storing his grain for i months. The value of $SC_i$ used in G-M's thesis is made up of two components. The first is a handling charge that the farmer must pay when he delivers his grain to the storage facility. This is a fixed cost and, as such, any change in handling costs would change the value of $SC_i$ by an equal amount, ceterus paribus, for all i.

The second component is the actual "rent" that the farmer pays to store his grain for a certain period of time. Obviously, this component of $SC_i$ increases as i becomes larger. If this charge is calculated on a per-month basis, then any equal across-the-board increase in this component would increase $SC_i$ by an amount equal to i times the increase per month. Since the 1976-77 monthly storage costs for corn and soybeans were in the \$.025 to \$.03 range after the initial three months, it would seem unlikely that they would change by much more than \$.01/month in any given marketing season. If we assume that the monthly charge had increased from given levels by \$.01/month, this would have yielded a maximum change in the $bep_i$ of \$.10 per bushel of corn in September and \$.10 per bushel of soybeans in August. Since the marginal increase in $SC_i$ is only \$.01/month, it would have had no effect on the actual optimal marketing period. Because $SC_i$ enters the $bep_i$ equation in an additive manner, only very large changes (relative to their 1976-77 levels) in storage and handling costs would have an effect on the optimal marketing time or on the ordering of marketing decisions.

# CHAPTER V. SUMMARY AND CONCLUSIONS

We have, in this thesis, looked at several alternatives for making decisions under uncertainty. We have investigated the use of Monte Carlo, portfolio, MOTAD, MALA, Bayesian and several game theory models. Because our objective was to simplify the calculations involved in arriving at some useable results, we decided to concentrate our empirical work on deriving a simplified procedure to calculate the Bayesian posteriors.

As mentioned throughout this paper, our work was based on a doctoral dissertation written by Hector Eduardo Gonzalez-Mendez. In his thesis, G-M developed a Bayesian decision model for corn and soybeans which, when coupled with some simple price forecasting models, generated values for expected marketing gains for each month of the marketing season. He did extensive work in the testing of the statistical underpinnings of the model and presented results for three forms of the Bayesian model. Our intentions in re-working G-M's model formulation were twofold. First, we wanted to derive a procedure for calculating the means of the DATA posterior distributions which would circumvent the tedious and time-consuming method outlined by G-M. Secondly, because we considered the distribution of states of nature to be continuous, we wanted a method of calculation which treated them as such.

We began by assuming that cash prices (states of nature) and forecasted cash prices were distributed as a bivariate normal. Given this assumption, we were able to show that the posterior distributions were normally distributed with easy-to-compute means and variances. We then proceeded to obtain values for the parameters involved in the mean and variance expressions. These took the form of some sample statistics as well as the slope coefficient of a simple linear regression. It was then possible to compile a listing of the expected gains to marketing for each combination of forecasting model, marketing period and commodity. This done, we compared our results with those obtained by G-M. Many of our numerical results were significantly different from G-M's due, most likely, to any of three reasons. First, there was a difference in the basic Bayesian model employed by each of us. Our model was a Bayesian DATA model with data priors whereas G-M's most comparable model was a DATA model with nondata priors. In other words, our prior probabilities were based on historical distributions of cash prices as opposed to his priors which were based on farmers' attitudes toward possible states of nature around the $bep_i$'s. The second source of difference was G-M's breaking of the distributions of the $S_j$ into discrete units. As just mentioned, our analysis treated this distribution as being continuous. The third source of variation involves our assumption of joint normality. While we did not justify this assumption statistically, our results seem to have at least given some support to the assumption.

Assuming that a given farmer would have marketed equal quantities of each commodity in the months indicated by both G-M's models and our models, his actual gain per bushel would have been greater under our formulation than under G-M's. While this is by no means an adequate comparison of the two different formulations, it does serve to point out that our procedure will yield plausible results with considerably less computation.

It would be useful to make a more thorough comparison of the two procedures. A method which evaluates the value of an experiment, as outlined by Halter and Dean (12, pp. 124-28), would seem to be a satisfactory method for carrying out this comparison. This method determines the expected value, in terms of money or utility, of the information used in computing the Bayesian strategy. Essentially, it begins by computing the expected payoff associated with the optimal Bayesian NONDATA strategy. The value of the additional information contained in the DATA model can then be computed as the difference between the expected payoff from the DATA model and the expected payoff from the NONDATA model. It would be possible to compute the expected payoffs under our procedure and under G-M's procedure to make a more accurate comparison.

We have, in conclusion, derived a procedure whereby Bayesian Decision Theory can be applied to the farm marketing decision with a minimum amount of computational effort. We are hopeful that this will lead to the successful derivation of further rules of thumb which can be used by the farm operator in his marketing decision. We do not presume to tell the farmer when to market his grain but, rather, we hope to present him with information which he can use in helping him decide the best course of action. He is still the person who assumes the risks and with whom the final decisions rest.

We would be remiss to close this thesis without taking at least one verbatim quote from G-M's dissertation. It involves the areas of research which should be pursued in order to further the work that each of us has done.

> "... we recognize the need for further research in at least two directions; 1) we have only considered the marketing decision problem at harvest time, it would be of interest to expand the model in a way that continuous revision of marketing decisions is possible ... 2) ... it would also be of interest to analyze many other alternative forecasting models and priors" [10, p. 160].

These are both problems that were alluded to earlier and we concur with G-M in stressing that further research be done in this area.

BIBLIOGRAPHY

1. Anderson, Jock R., Dillon, John L. and Hardaker, J. Brian. Agricultural Decision Analysis. Ames, Iowa: The Iowa State University Press, 1977.

2. Borch, Karl. "The Rationale of the Mean-Standard Deviation Analysis: Comment." American Economic Review 64 (June 1974):428-30.

3. Boussard, J. M. "A Model of the Behavior of Farmers and Its Application to Agricultural Policies." European Economic Review 2 (Summer 1971):436-61.

4. Boussard, J. M., and Petit, M. "Representation of Farmers' Behavior Under Uncertainty With a Focus Loss Constraint." Journal of Farm Economics 49 (November 1967):869-80.

5. Davies, O. L., and Pearson, E. S. "Methods of Estimating From Samples of Population Standard Deviation." Journal of the Royal Statistical Society (Supplement) 1 (1934):76-93.

6. Dillon, John L., and Heady, Earl O. "Theories of Choice in Relation to Farmer Decisions." Iowa Agricultural and Home Economics Experiment Station, Research Bulletin 485, 1960.

7. Fisher, R. A. "A Mathematical Examination of the Methods of Determining the Accuracy of an Observation by the Mean Error, and by the Mean Square Error." Royal Astronomical Society (Monthly Notes) 80 (1920):758-69.

8. Geary, R. C. "Moments of the Ratio of the Mean Deviation to the Standard Deviation for Normal Samples." Biometrika 28 (1936):295.

9. Goldman, Steven M. "Portfolio Choice and Flexibility: The Precautionary Motive." Journal of Monetary Economics 4 (April 1978):263-79.

10. Gonzalez-Mendez, Hector Eduardo. "Grain Marketing Gains in Iowa and the Use of Price Forecasting Models - A Bayesian Decision Approach." Ph.D. dissertation. Iowa State University, Ames, Iowa, 1977.

11. Gonzalez-Mendez, H. E., and Ladd, George W. "Bayesian Grain Storage Decisions and Use of Price Forecasting Models." Unpublished Journal Paper. Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, 1978.

12. Halter, A. N., and Dean, G. W. Decisions Under Uncertainty. Cincinnati, Ohio: South-Western Publishing Company, 1971.

13. Hazell, P. B. R. "A Linear Alternative to Quadratic and Semi-variance Programming for Farm Planning Under Uncertainty." American Journal of Agricultural Economics 53 (February 1971): 53-62.

14. Hurwicz, L. "Optimality Criteria for Decision Making Under Ignorance." Cowles Commission Discussion Paper, Statistics, No. 350, 1951. (Mimeographed.)

15. Kennedy, J. O. S., and Francisco, E. M. "On the Formulation of Risk Constraints for Linear Programming." Journal of Agricultural Economics 25 (May 1974):129-44.

16. Kolstoe, Ralph H. Introduction to Statistics for the Behavioral Science. Homewood, Illinois: The Dorsey Press, 1969.

17. Koopmans, Tjalling C., ed. Activity Analysis of Production and Allocation. Proceedings of a Conference. New York: John Wiley and Sons, Inc., 1951.

18. Ladd, George W. Taken from Economics 648 Lecture Notes. Iowa State University, Ames, Iowa.

19. Ladd, George W., and Williams, Joseph E. "Application of Net Energy System and Bayesian Decision Theory for Determination of Cattle Rations and Rates of Gain." Iowa Agriculture and Home Economics Experiment Station, Journal Paper Number J-8636, 1978.

20. Levy, Haim. "The Rationale of the Mean-Standard Deviation Analysis: Comment." American Economic Review 64 (June 1974): 428-30.

21. Markowitz, Harry M. Portfolio Selection: Efficient Diversification of Investments. New York: John Wiley and Sons, Inc., 1959.

22. Mood, Alexander M., Graybill, F. A., and Boes, D. C. Introduction to the Theory of Statistics. New York, New York: McGraw-Hill Book Company, 1974.

23. Morrison, Donald F. Multivariate Statistical Methods. New York: McGraw Hill Book Company, 1976.

24. Mossin, J. "Aspects of Rational Insurance Purchasing." Journal of Political Economy 76 (July-August 1968):553-68.

25. Savage, L. J. "The Theory of Statistical Decision." Journal of the American Statistical Association 46 (March 1951): 55-57.

26. Snedecor, George W. and Cochran, W. G. Statistical Methods. Ames, Iowa: Iowa State University Press, 1973.

27. Thrall, R. M., Coombs, C. H., and Davis, R. L., eds. Decision Processes. New York: John Wiley and Sons, Inc., 1957.

28. Tobin, J. "The Theory of Portfolio Selection." In the Theory of Interest Rates, pp. 18-33. Edited by F. Hahn and F. P. R. Brechling. New York: St. Martin's Press, 1965.

29. Tobin, J. "Liquidity Preference as Behavior Toward Risk." Review of Economic Studies 25 (February 1958):65-86.

30. U. S. Department of Agriculture. Economic Research Service. Agricultural Finance Statistics. Annual Series. Farm Mortgage Interest Rates. Washington, D. C.: U. S. Government Printing Office, 1977.

31. Van Horne, James C. Financial Management and Policy. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1977.

32. von Neumann, John and Morgenstern, Oskar. Theory of Games and Economic Behavior. Princeton, New Jersey: Princeton University Press, 1944.

33. Wald, A. Statistical Decision Functions. New York: John Wiley and Sons, Inc., 1950.

34. Walker, Odell L., Heady, Earl O., Tweeten, Luther G., and Pesek, John T. "Application of Game Theory Models to Decisions on Farm Practices and Resource Use." Iowa Agriculture and Home Economics Experiment Station, Research Bulletin 488, 1960.

# APPENDIX

## Reliability of the Taylor Expansion

It is necessary to make note of the entire concept of maximizing utility or, more correctly, maximizing expected utility. The idea of expected utility enters here on the basis of the decision maker's confrontation with actions which will yield returns in the future; i.e., the individual cannot maximize actual utility because he is faced with an uncertain return.

Measurement of expected utility is a problem all of its own. According to von Neumann and Morgenstern [32], the expected utility of a distribution of uncertain returns can be determined relative to the actual utility obtained from a return that is received with certainty. This can be illustrated as follows. Assume three certain events, A, B and C. Assume also that event A is preferred to event B and that event B is preferred to event C or, notationally,

$$A > B > C$$

If we now introduce an uncertain event composed of obtaining A with probability p and obtaining C with probability (1-p), the von Neumann-Morgenstern axioms state that there exists a value of $p (0 \le p \le 1)$ such that the individual will be indifferent between the uncertain prospect composed of A and C and the certain prospect, B. Therefore,

$$\text{(A.1)} \quad E[u(A \text{ or } C, p)] = pu(A) + (1-p)u(C)$$
$$= u(B)$$

This, then, defines the expected utility of the uncertain event, given p. Note also that, knowing p and arbitrarily assigning values for two of the certain events, the utility of the other event can be determined. This process can, of course, be continued to include as many events or combinations of events as desired. The most important point to be made here, however, is to simply state what expected utility is and to show that it is a measurable entity.

If states of nature are distributed as continuous variables (as opposed to the discrete case cited above), then estimation of the expected utility becomes somewhat more difficult. Assume that the states of nature, s, are distributed continuously as L(s). Assume also that the payoffs of action $a_i$, the $x_i$, are a continuous function of s, distributed as $f(x_i)$. Given these assumptions, the probability of obtaining some $x_i$, $x_{i\alpha} \le x_i \le x_{i\beta}$, can be calculated by

$$\text{(A.2)}\quad P(x_{i\alpha} \le x_i \le x_{i\beta}) = \int_{\alpha}^{\beta} f(x_i)\, dx_i$$
$$= P(s_{\alpha} \le s \le s_{\beta})$$
$$= \int_{\alpha}^{\beta} L(s)\, ds$$

In practice, most researchers (relying on the von Neumann-Morgenstern axioms for support) assume that a specific utility function exists in the form

$$\text{(A.3)}\quad u = u(x)$$

or, given that action i has occurred, $u_i = u(x_i)$. By appealing to the mathematical derivation of the expected value of a continuous variable, we can now obtain an expression for the expected utility of action i:

$$\text{(A.4)} \quad E(u_i) = \int_{a_i}^{b_i} u(x_i)\, f(x_i)\, dx_i$$

where $a_i$ and $b_i$ define the range of variation in $x_i$.

In many cases

$$\int_{a_i}^{b_i} f(x_i)\, dx_i$$

can be integrated while

$$\int_{a_i}^{b_i} u(x_i)\, f\,(x_i)\, dx_i$$

cannot. It is because of this dilemma that researchers have turned to alternate methods of deriving the value of $E(u_i)$ [18].

The most widely accepted alternative method of obtaining expected utility involves expansion of the utility function using the Taylor series. The following derivation is quoted from <u>Decisions Under Uncertainty</u> by Albert N. Halter and Gerald W. Dean [12].

The function u(x) can be expanded to a function in powers of (x-c) where x is a random variable and c is a fixed value. In particular, the Taylor series expansion of u(x) is:

$$\text{(1)} \quad u(x) = u(c) + (x-c)\,\frac{du(c)}{dx} + \frac{1}{2}\,(x-c)^2\,\frac{d^2u(c)}{dx^2}$$

$$+ \frac{1}{3!}\,(x-c)^3\,\frac{d^3u(c)}{dx^3} + \frac{1}{4!}\,(x-c)^4\,\frac{d^4u(c)}{dx^4} + \ldots.$$

Letting c = E(x), expected gain for any action, we obtain

$$(2)\quad u(x) = u[E(x)] + [x-E(x)]\,\frac{du[E(x)]}{dx} + \frac{1}{2}\,[x-E(x)]^2\,\frac{d^2u[E(x)]}{dx^2}$$

$$+ \frac{1}{3!}\,[x-E(x)]^3\,\frac{d^3u[E(x)]}{dx^3}$$

$$+ \frac{1}{4!}\,[x-E(x)]^4\,\frac{d^4u[E(x)]}{dx^4} + \ldots.$$

Taking the expectation of each side of the equation, we obtain the expected utility of action a:

$$(3)\quad u(a) = Eu(x) = u[E(x)] + \frac{1}{2}\,\sigma^2\,\frac{d^2u[E(x)]}{dx^2}$$

$$+ \frac{1}{3!}\,g_1\,\frac{d^3u[E(x)]}{dx^3} + \frac{1}{4!}\,g_2\,\frac{d^4u[E(x)]}{dx^4} + \ldots.$$

where

the expectation of the constant $E(x) = E(x)$,

the expectation of $x-E(x) = 0$

the expectation of $[x-E(x)]^2 = \sigma^2$ i.e., the variance of the distribution of x,

the expectation of $[x-E(x)]^3 = g_1$, i.e., the skewness of the distribution of x,

the expectation of $[x-E(x)]^4 = g_2$, i.e., the kurtosis of the distribution of x. ([12], pp. 100-101).

This is the Taylor series expansion <u>about the mean</u>.

The problem with this procedure comes about because of the remainder whose existence is indicated by the dots in (3). (A remainder exists because the Taylor series expansion is only an approximation to the true value of the function.) In general, the Taylor expansion can be written as

$$\text{(A.5)} \quad f(x) = f(a) + f'(a)\,(x-a) + \frac{f''(a)}{2!}\,(x-a)^2$$

$$+ \ldots. + \frac{f^{(n)}(a)}{n!}\,(x-a)^n + \frac{f^{(n+1)}(x_1)}{(n+1)!}\,(x-a)^{n+1}$$

For some $x_1$, such that $a < x_1 < x$ if $x > a$

and $a > x_1 > x$ if $x < a$

The final term of (A.5) is the remainder. In terms of the utility function considered earlier, (A.5) can be written as

$$\text{(A.6)} \quad u(x) = u[E(x)] + [x-E(x)]\,\frac{du[E(x)]}{dx} + \frac{1}{2}\,[x-E(x)]^2\,\frac{d^2u[E(x)]}{dx^2}$$

$$+ \frac{1}{3!}[x-E(x)]^3\,\frac{d^3u[E(x)]}{dx^3} + \frac{1}{4!}[x-E(x)]^4\,\frac{d^4u[E(x)]}{dx^4}$$

$$+ \frac{1}{5!}[x-E(x)]^5\,\frac{d^5u(x_1)}{dx^5}$$

for some $x_1$ such that $E(x) < x_1 < x$ if $x > E(x)$

and $E(x) > x_1 > x$ if $x < E(x)$

Taking the expectation of (A.6) and defining the last term to be $R_4$ yields

$$\text{(A.7)} \quad Eu(x) = u[E(x)] + \frac{1}{2}\,\sigma^2 \frac{d^2u[E(x)]}{dx^2} + \frac{1}{3!}g_1 \frac{d^3u[E(x)]}{dx^3}$$

$$+ \frac{1}{4!}\; g_2 \frac{d^4u[E(x)]}{dx^4} + R_4$$

It is obvious that if the utility function is of degree four or less (where the degree is an integer value), then

$$\text{(A.8)} \quad R_4 = \frac{1}{5!}\; g_3 \frac{d^5\; u[x_1]}{dx^5} = 0$$

$$\text{since } \frac{d^5u}{dx^5} = 0.$$

It is also apparent that if all moments of the distribution of payoffs beyond $g_2$ are equal to zero, then $R_4 = 0$ regardless of the degree of the utility function. Obviously, there is nothing magic about using four terms in the expansion. The same arguments concerning $R_n$ are valid for utility functions of degree n or less and distributions of payoffs with moments beyond $g_{n-2}$ all equal to zero.

In most empirical work every term to the right of $\frac{1}{2}\,\sigma^2 \frac{d^2u[E(x)]}{dx^2}$ is assumed to be zero in order to simplify the problem somewhat. If this is the case, then utility functions of degree 3 or greater coupled with payoff distributions possessing some $g_i$ ($i = 1, 2, \ldots$) which are nonzero (assuming nonzero $g_i$ correspond to nonzero $d^iu/dx^i$) yield $R_n \neq 0$. In the vast majority of cases it is assumed that the payoffs are normally distributed. In these cases,

$g_i \neq 0$ when i is even and $g_i = 0$ when i is odd. It is then possible to have $R_n = 0$ even with a cubic utility function. Since $g_1 = 0$, then $\frac{1}{3!}\, g_1 \frac{d^3u(x_1)}{dx^3} = 0$ regardless of the fact that $\frac{d^3u}{dx^3} \neq 0$ [18].

It is apparent that unless the assumptions concerning the degree of the utility function and the moments of the payoff distribution are met, the Taylor series expansion is only an approximation to the actual expected utility. If these approximations are sufficiently in error, it is quite possible that the indifference curves derived from them may yield nonoptimal decisions.