Risk

DEPARTMENT OF AGRICULTURAL ECONOMICS

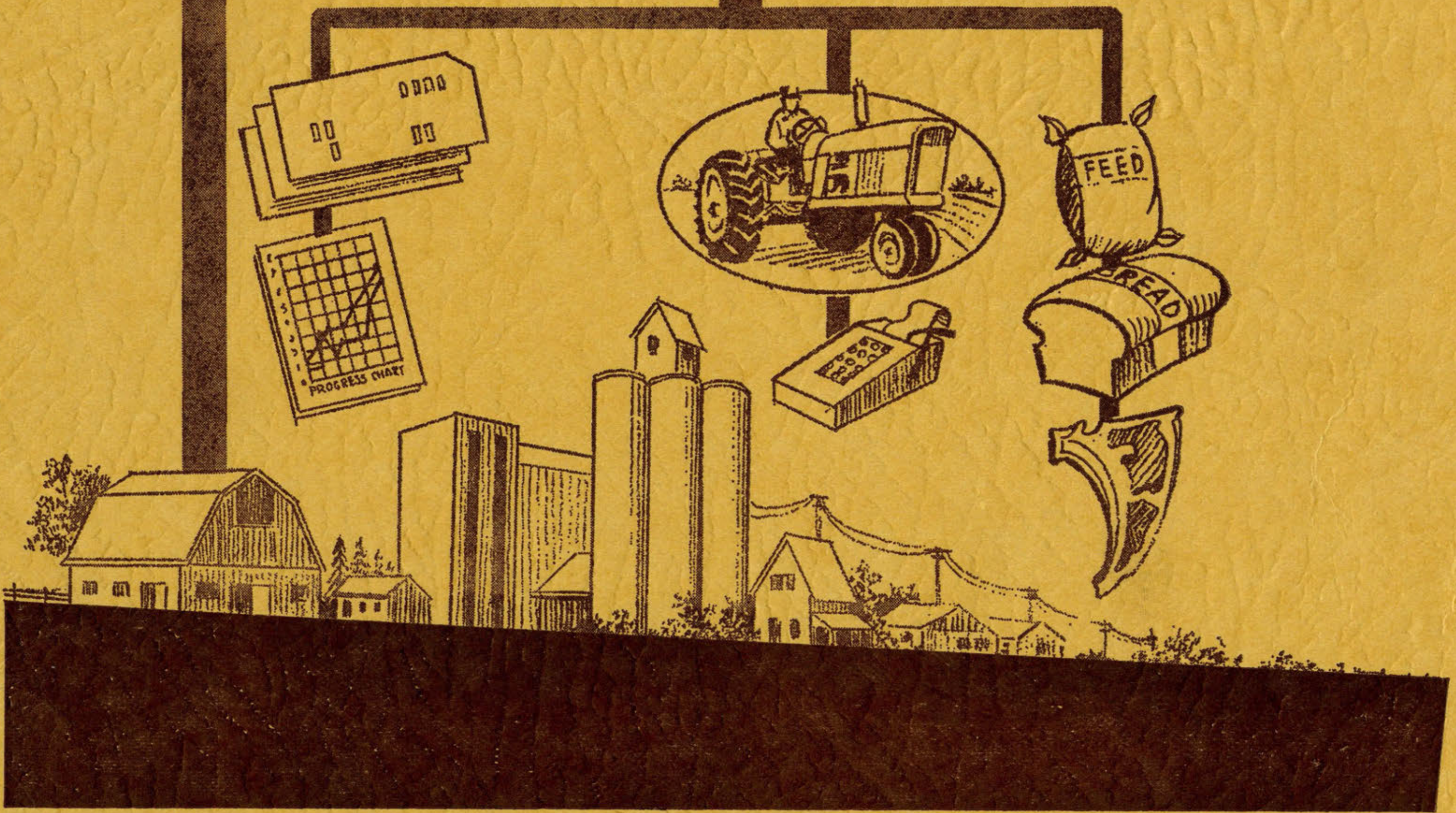
staff paper

A CRITICAL EXAMINATION
OF THE THEORY AND PRACTICE
OF QUADRATIC RISK PROGRAMMING

By

F.W. Barney

1983-#5



COLLEGE OF AGRICULTURE
UNIVERSITY OF NEBRASKA-LINCOLN, NEBRASKA

A CRITICAL EXAMINATION OF THE THEORY AND PRACTICE OF QUADRATIC RISK PROGRAMMING*

by

F.W. Barney

Instructor

Agricultural Economics Department

University of Nebraska-Lincoln

Lincoln, Nebraska

The Author wishes to acknowledge Glenn Helmer's (UN-L-Agricultural Economics) assistance and Pam Holmes' typing; the remaining mistakes and misdirection are the property of the author.

* Contributed paper read at the 1983 American Agricultural Economics Association Annual Meeting.

ABSTRACT

The theory and practice of Quadratic Risk Programming is critically reviewed and found to contain three major weaknesses: (1) There is no necessary connection between QP and expected utility. Thus it is not clear in what sense QP models can aid decision making under uncertainty or our understanding of it. (2) There is no theory telling the model builder how to operationalize a QP model. Since the estimated optimal set of activities in a QP model is a function of how the model was operationalized, this lack of theory is a crucial problem to those who wish to interpret and/or compare the results of various QP models. (3) The theory of QP gives us no guide as to the relationship of the model's estimated optimal set of activities and the set of activities that would result if the model was literally true and correctly operationalized. Given the above points, it is concluded that QP and its linear approximation MOTAD produce results that bare an unknown relationship to their underlying theory.

INTRODUCTION

Markowitz observed that the profit maximization assumption could not explain diversification, but the possession of a mean variance utility function by a decision maker could. Freund building on the work of Markowitz observed that linear programming tracted farm decisions poorly; but if the objective function of the linear program was modified by adding a quadratic term then farm planning behavior was replicated very well. Hazell showed that Freund's approach could be approximated by a procedure that is known as MOTAD. These three have collectively provided the rational for a considerable amount of applied risk analysis.

In this paper, the theory of Quadratic Risk Programming, in the sense of Freund, is critically reviewed. This review is from the point of view of the applied economist. By the point of view of the applied economist we mean: (1) The connection of QP with rational decision making under uncertainty is of interest; (2) How QP models are operationalized is of interest; and (3) the interpretation of QP models is of interest. The review of QP from this perspective is undertaken with three purposes in mind: (1) to decide in what sense QP is an applied economic tool; (2) to shed some light on the current discussion between Buccola, Johnson and Boehlje about QP and MOTAD, and (3) to discuss some methodological questions that are common to many risk programming methods.

This paper has five more sections. In the next section we introduce definitions and notation. In the second the connection between QP and expected utility is explored. In the third the data needs of a QP model are discussed. The fourth discusses the sensitivity of QP models to specification error. The fifth section concludes the paper by drawing a parallel between QP and ridge regression.

DEFINITIONS AND NOTATION

Consider the following LP problems:

Maximize $Z = CX$

Subject to $AX \leq b$

$X \geq 0$

Where C is a vector of net per unit revenues of a set of production processes available to the farm, A denotes the technology matrix, X the choice variables and b the available resources. Freund noted that the application of LP to farm planning required C to be known. This requirement of LP doesn't consider price risk and thus is the source of LP's inability to track farm planning decisions.

To incorporate price risk into a programming model, Freund made the following assumptions:

(1) $C \sim N(\mu_i, \sigma i_i^2)$ for all i. That is the conjectured per unit net revenue of the i^{th} process in normally distributed with a mean of μ_i and a variation of σ_i.

(2) The decision maker is an expected utility maximizer with utility function, $u(r) = 1 - e^{-ar}$, in net reveue, r. $r = CX$ and $a > 0$, which represents the decision maker's aversion to risk. With

$u'(r) = ae^{-ar} > 0$ marginal utility of r is positive; and

$u''(r) = -a^2e^{-r} < 0$ global risk aversion.

What Freund was able to show is that the maximization of

$$Z = CX - \frac{aX^T \Sigma X}{2}$$

such that $AX \leq b$

$X \geq 0$

where Σ is the variance covariance matrix of the vector of net returns when all choice variables are at unit level, maximized the decision maker's expected utility. Freund's argument about the source of the weakness of LP farm planning models while not complete, was a clear improvement. His solution was ingenious. His modification of LP clearly fit his data better than a LP model.

QP AND EXPECTED UTILITY

It is now known that QP is consistent with expected utility maximization in only two cases (Tobin): (1) the utility function is a polynomial in r and the net returns variable is normally distributed or (2) the decision maker has a quadratic utility function. Despite the many times normality has been invoked in this context the economic justification for the assumption of normal distributed net returns is zero. Bankruptcy laws assure us that all downside risk is finite. Since an infinite downside risk precluded by these laws, no producer will have a subjective assessment of net returns that is normal. Unless the decision makers subjective distribution of returns is exactly normal, it is not possible to express all higher order moments of the utility function as functions of the mean and variance of r. If this cannot be done the maximization of $C^TX - \frac{AX^T\Sigma X}{2}$ no longer maxmimizes the decision makers expected utility.

Quadratic utility functions can be expressed as mean variance utility functions and thus are thought to be consistent with expected utility maximization. The matter is more complex than is commonly realized. Consider the following Bernoulli experiments: (1) P is the probability of success, 1-P is that of failure, success pays out X_1, failure 0. (2) P is the probability of success, 1-P is that of failure. Success pays out X_2, failure 0. Assume $X_1 > X_2$. Then the dominance axiom implies experiment two is preferred to one. Furthermore there is no $X_2 > X_1$ such that one is preferred to two. The above argument due to Karl Borsh shows that in general mean variance indifference curves and their quadratic utility functions are not in general consistent with expected utiity theory. Only if returns are normally distributed is quadratic utility consistent with expected utility maximization.

The example raises the possibility of EV analysis without the assumption of expected utility maximization. This possibility divides into: (1) the decision maker's preferences are exactly described by a mean variance utility function; alternatively (2) such a function is a good approximation of the decision makers true preferences.

With reference to the first case, observe that it is formally correct to say that a mean variance utility function and Freund's assumptions imply that QP is consistent with expected utility maximization. The problem with this view is that there is not the slightest bit of evidence in theory for the use of σ^2 as a measure of risk. Indeed the use of σ^2 as a measure of risk is counter to common usage of the term. When risk is used, it is in connection with adversity (i.e. windfall losses). Windfall losses are something that we take insurance out against. Windfall gains invoke thanks to the Almighty, not sighs that we had sense enough to protect ourselves against windfall increases in income. Using variance to measure risk implies that responses to unexpected changes in income are symmetrical. This view, to say the least, runs counter to the common use of the term risk. This is not to say that those who wish to measure risk by σ^2 are wrong. It is only to say that it is not obvious why everyone would behave as if they possess a mean variance utility function.

On an analytical level, there are several other measures related to variance that are reasonable proxies for the risk associated with a given plan. Consider the following measure of risk of the i^{th} farm plan where $p_i(r<\lambda) = f_i(\lambda)$ and $f(\lambda)$ is the density function of $F_i(\lambda)$.

$R_i = {}_{-}\int^{\lambda} (c-x)^{\alpha} f(x)dx$, R_i defines a general risk function. For $\alpha = 0$ we have $R_i = P_i(X<\lambda)$, target rate risk. For $\alpha=2$ and $\lambda=c$ we have $R_i = {}_{-\infty}\int^{c}(c-r_i)^2F(x)dx$ semivariance risk. For $\alpha=2$, $c=E(r)$ and $\lambda=\infty$, $R_i = {}_{-\infty}\int^{\infty} (E(r)-x)^2 f(x)dx = \sigma_i^2$;

variance risk. What has been shown is that σ_i^2 is a specific case of a general risk function. By deducing three popular risk measures from R_i, we have shown that these risk measures are related. Which values of λ, c and α yield the "best" measure of risk is not a settled matter.

In a more fundamental sense, what justification do we have for believing that risk is one dimensional? In a farm planning context why do we not consider the following two dimensional measure of risk; the first component is the probability that a farm plan will lead to the cost of acquiring capital next year rising, the second component being the probability that the same plan will lead to loss of the farm. I submit this view of risk has more to commend it than the one dimensional variance view. In the absence of a solid theory from any of the social sciences justifying variance as a measure of risk, it appears reasonable to keep an open mind about the merits of lack there of risk models that use variance as a measure of risk. It is true that when applied they may yield useful results; but these results have not been connected in a firm manner to existing economic theory.

What is at issue is the nature of risk. We seem to agree that risk invlves the possible states of the world, say S_i at i = 1 . . m. It also involves consequences to the decision maker, C_{jk} at J = 1 . . . n, read the consequences of plan K given state J prevailed. Observe C_{jk} need not be one dimensional. Finally there is agreement that it involves the probability of the states of the world, P_i = 1 . . . m. Thus we all agree that (S_i, C_{jk}, P_i) are components of risk. However, we have no agreed upon theory that allows us to aggregate (S_i, C_{jk}, P_i) into a single risk measure for plan K. The use of variance as a measure of risk is one of an infinite number of ways to aggregate the components of risk. To those who extol the merits of mean variance analysis, it is reasonable to ask what basis other than tradition and ease of computation do you have for using variance as an indicator of risk?

While the above argument is correct, some find it to be pedantic. Levy and Markowitz have argued that what is of interest to applied economists is sound approximation not analytic exactness and by this standard, mean variance analysis performs quite well. Even if we accept the argument as stated, it is spurious. It matters not to applied economists that the two variable utility function is a sound approximation of the N variable; but rather that the otimal plan is close to the extimated plan. Close in utility space which is one dimensional cannot bare any clear relationship to close in activity space which is N dimensional.

Let $U(\underline{X}c)$ equal the utility of the bundle chosen and let $\hat{U}(\underline{X}p)$ be the utility of the predicted bundle. Close in utility space means for $\varepsilon > 0$ and small, we have that $|\hat{U}(\underline{X}p) - U(\underline{X}c)| < \varepsilon$. Since $\underline{X}p$ and $\underline{X}c$ are elements of R^n, it is not clear what $|\hat{U}(\underline{X}p) - U(\underline{X}c)| < \varepsilon$ implies about $\underline{X}p - \underline{X}c$. Thus the approximation view of mean variance analysis does not produce an approximation that is of use to those economists interested in resource allocation.

To those of us that have grown up with the computer and have studied numerical analysis, it is self evidently true that a quadratic function can be constructed to provide a good local approximation to any function. The problem with self evident truths is that they flourish only in unexplored regions of the mind! Let $T(U) = A + BU$ where A and B are elements of R^1_+. Given $\hat{U}(\underline{X}p) - U(Xc) < \varepsilon$ then $T(U(Xp)) - T(\hat{U}(Xc)) = B(U(Xp)) - B(\hat{U}(Xc)) < B\varepsilon$. Since the utility functions we work with are unique up to an affine transformation, the above argument shows that it is impossible to speak of close in the Euclidean norm between observed and predicted utilities.

It is possible to define a metric such that the mean variance utility function may be thought of as a close approximation to the true utility function, but why bother? The solution to this psychological problem bares an unclear relation to the economic problem, the allocation of resources.

DATA REQUIREMENTS OF A QP MODEL

Freund's original formulation of expected utility argued effectively for the empirical improvements of QP over LP. In operationalizing QP, the analyst is compelled to estimate a and Σ. Correctly done, this procedure is far more subtle than is commonly realized. Current practice involves detrending data and using the resulting data to estimate the risk aversion coefficient, a, and Σ. This procedure assumes that: (1) the analyst has the appropriate data to estimate Σ and (2) that he knows the appropriate weight to give to each observation. Neither of these assumptions appears reasonable.

To estimate Σ, the analyst must use the decision makers subjective estimates of the variability associated with each plan. To see that market data can not be used, consider the following thought experiment. Consider a hundred almost identical corn farmers with 100 identical corn farms. They have identical feelings about risk in that if they were given $100 to gamble, they would all make the same bets. They differ only in their ability to grow corn. Some are excellent farmers, others are not and the bulk are somewhere in the middle. Supposse that after considering the same information 20 decide to participate in PIK and 80 decide against participation. In my example, all farmers possess the same risk preferences but different levels of skills. By skills I mean the ability to effect both the mean level of return and its variability. Being aware of their own skills, they are able to correctly assess the risk that they face in the market. Thus, the same risk performance and different skill levels produce a 20/80 split.

One can just as well imagine the case where all farms are equally skilled and a 20/80 split is a function of different risk preferences. By using past market data to estimate a and Σ we ignore the fact that the risk a farmer faces is a function of his skill and the randomness in the system. Thus, to operationalize a QP model requires us to use data that reflects only the farms perceived randomness. This requires that we elicit his subjective distribution of net returns. Failure to do so produces results that are an unspecified function of risk preferences and skill differences.

Besides ascertaining the data used to form expectations, the QP modeler must be able to say what weight the decision maker gives to each observation before he has a data base from which we can construct a QP model that will be an accurate representation of his risk preferences. The custom in these matters appears to be that of treating each observation equally. While this has a fine democratic ring to it, I rather doubt that any decision maker weighs the distant past and the close past on the same scale. The equal weighting scheme has some nice statistical properties but it does not possess a theoretical justification. To estimate Σ, we are required to view the world through the decision maker's eyes. Thus we are required to have some theory, other than the axiom of convenience, of what constitutes data.

We would like to think that the above is unjustified; after all Tintner's work seems to imply that we can make reasonable estimates of the noise in any series by a relatively simple approach based on the assumption that all values are equally likely. This argument is fundamentally correct, however its use is suspect. If QP model is a correct characterization of the decision makers risk preferences and objective constraints, the solutions using an equal weighting scheme and the decision maker's weighting scheme can be different. How different depends upon the problem at hand.

In summary, the logic of choice implies that decisions are a function of the decision makers expectations. If we are constructing a QP model that we hope some decision maker will find valuable, we are required to estimate on the same set of data with the same weights as he would. Failure to do otherwise may produce results that differ from what we would obtain by using the correctly defined data. In the end, we have to use imperfect data to try to solve the economic problems we face. What this section argues is that there is considerable reason for feeling that the results of a QP model are sensitive to the analyst's data definitions. Our inability to clearly define the data we should be using in our QP model makes their sensitivity to specification error a problem of considerable magnitude.

SENSITIVITY OF QP MODELS TO ESTIMATION ERRORS

The continued use of QP models implies that the designers and users of the models find the meaningfulness of their results sufficient to justify the limitations of their methodology. This conclusion reflects a belief that the estimated set of optimal activities is close to the optimal set of activities. If this is the case, they are right - objections to the theory and practice of QP have few practical implications. What follows is a two variable argument to the effect that any estimation error produces a set of predicted activities that bares an unknown relationship to the set of activities that would result if QP was a literal representation of the decision makers preferences and the constraints on his behavior.

To understand the implications of Freund's modification to LP, consider the following two variable example:

$$Z = C_1X_1 + C_2X_2 - \left[\frac{a}{2}X_1^2\sigma_1^2 + aX_1X_2{}_{12} + \frac{a}{2}X_2^2\sigma_2^2\right]$$

where σ_{12} denotes the covariance between the net revenue when processes 1, 2 are run at unit level. For $Z = Z_*$, the total differential of Z_* is

$$0 = (C_1 - aX_1\sigma_1^2 - aX_2\sigma_{12})dX_1 + (C_2 - aX_2\sigma_2^2 - aX_1\sigma_{12})dX_2$$

or $$\frac{dX_2}{dX_1} = -\frac{(C_1 - aX_1\sigma_1^2 - aX_2\sigma_{12})}{(C_2 - aX_2\sigma_2^2 - aX_1\sigma_{12})}$$

By inspection of the two variable case it is observed that attaching $-\frac{a}{2}X^T\Sigma X$ to the objective function has the effect of altering both the gradient of the objective function and its contours. Given these alterations, it is not surprising that QP tends to generate a different set of optimal activities than LP. This is a point of considerable practical importance. Since the difference between the LP and QP solutions is a function of a, and Σ the correct estimation of a and Σ is necessary if we are to produce a QP model that is an improvement over LP and is consistent with a decision makers preferences. Failure to correctly estimate a and Σ results in producing an estimated plan that has no clear relationship to the structure of the model. In this context I find it interesting to note that in Freund's original article $a = \frac{1}{1250}$ was chosen because $\frac{1}{2500}$ was too small!

For any given constraint structure, the solution to a QP problem is a function of a and Σ, neither of which are known. Thus, both must be estimated. In applied QP analysis, a given set of data is used to estimate a and Σ; $\hat{a}$ and $\hat{\Sigma}$ are then placed in the model which is to be maximized. The resulting set of optimal activities is used as if no estimation was involved in the process. To make use of results based upon a and Σ requires some theory defining the distribution of our estimates of a and Σ and the distribution of estimated optimal activities. Furthermore, we require a knowledge of the link between the estimated optimal activities and the set of activities that would result if we used a and Σ. Unfortunately, for those that are interested in QP, there is no theoretical justification for the assumption that the estimated results will be "near" what would result from using the parameter of the model. The estimated set of optimal activities is a function of: (1) the length of the time series used, (2) the weight given to each observation and (3) the constraint structure. For one set of constraints the optimal solution can be insensitive to estimation errors; for another the reverse can be true. There is no way in theory to decide which is the case! Thus, applied QP is similar to ridge regression; both can yield results of interest but in neither case can we say what relationship the estimated values have to the corresponding population parameters. Since there are no clear links between the parameters of a QP model and our estimates, the interpretation of QP models is far from clear.

CONCLUSIONS

Where does this leave us? When we view QP as modified LP, we see that the results we obtain are a function of the risk aversion parameter and the variance covariance matrix of returns when all activities are run at a unit level. Since we must estimate both parameters, a QP model yields an estimated optimal set of activities. In view of the fact that QP has no theory connecting its estimated optimal set of activities to the optimal bundle, the results of QP models and their linear approximation MOTAD bare no clear relationship to their underlying theory. Thus such models must be thought of as forecasting models.

The lack of normality of returns lessens the connection of QP models with expected utility models. The lack of any theory justifying variance as measures of risk undermines completely the connection of QP with expected utility. Thus, there is no reason why the behavior of farm managers should follow the QP model.

In view of the before mentioned arguments, I find the current discussion about the relative merits of MOTAD and QP long on mathematical assumptions and short on economic reality.

REFERENCES

Borch, K. "A Note on Uncertainty and Indifference Curves." Review Economic Studies: 36(1969):1-4.

Buccola, Steven T. "Minimizing Mean Absolute Deviations to Exactly Solve Expected Utility Problems: Comment." American Journal of Agricultural Economics: 64(1982):789-791.

Freund, R.J. "The Introduction of Risk Into a Programming Model." Econometrica: 24(1956):253-63.

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

Johnson, Donald and Boehlje, Michael. "Minimizing Mean Absolute Deviations to Exactly Solve Expected Utility Problems: Reply." American Journal of Agricultural Economics: 64(1982):792-793.

Levy, H. and Markowitz, H.W. "Approximating Expected Utility by a Function of Mean and Variance." American Economic Review: 69(1979):308-316.

Markowitz, H. "Portfolio Selection." Journal of Finance: 7(1952): No. 3:82-92.

Tintner, G. The Variate Difference Method. Bloomington, Illinois: Principia Press, 1940.

Tobin, J. "Comment on Borch and Feldstein." Review of Economic Studies: 30(1969):13-14.