GUIDELINES FOR APPLIED AGRICULTURAL RESEARCH: DESIGNING, REPORTING AND INTERPRETING EXPERIMENTS

J. R. Anderson*

Commentary is made on the design, reporting, and economic interpretation of field experiments, and attention is given to Bayesian statistics as a potentially useful aid in planning and assessing research. This article complements the discussions of Lloyd (1958) and Dillon (1966) in this *Review*.

1 INTRODUCTION

The recent and continuing decline of several agricultural industries in Australia may be in part responsible for the current questioning of the nature and orientation of much of our agricultural research programme. An important feature of this questioning attitude is the increasing interest being given to the economic implications of experimental work conducted in agricultural and veterinary science.

Some of the interest in planning experiments with a view to subsequent economic interpretation may have been kindled by the evangelical writings of such people as Heady in the U.S.A. [16, 17, 18], Wragg in the U.K. (see e.g. [33]) and Dillon [9, 10, 18] and Lloyd [22] in Australia. The present article was drafted in response to such developing interest on the part of agricultural researchers and is designed to complement the earlier discussions of Dillon [9] and Lloyd [22].

For the purpose of this review a primary distinction is made between research and experimental work intended to influence decisions about the management of an economic system and "other" research. For the former it is possible to place some estimate on the economic consequences of making wrong decisions (i.e. estimate loss functions) whereas for the latter, loss functions cannot be specified readily in economic terms. This article relates only to the former category of research which for semantic convenience might be labelled applied research. Categories of research with a more "basic" orientation in which, for example, questions such as "Why?" and "Where Next?"

---

* University of New England.
The critical comments by John L. Dillon, Bruce Standen, and David Trebeck were useful in revising this paper which, with the participation of Alan G. Lloyd, was presented to a Course on Pastoral Experiments for Field Research Workers held at the State Research Farm, Werribee, February, 1971.

are posed are not so amenable to economic analysis, and are conveniently avoided here. Dillon [9] has usefully classified the experiments of applied research (or decision-oriented experiments) into "Yes/No" experiments and "How Much?" experiments. These two categories are discussed in sections 2 and 3 respectively.

2 YES/NO EXPERIMENTS

The design requirements for "Yes/No" experiments are simple and traditional. Typically, these experiments will essentially involve a comparison of treatment means, and designs commonly used include randomised blocks, Latin squares or more complex designs such as crossover, split plot, and lattice designs. The types of question asked in these experiments would be, for example, "Is this new drug superior to the current standard?"; "Do these strains differ from one another in productivity in this environment?"; "Is this new grazing system really any different from the traditional district system?".

The traditional practice of interpreting "Yes/No" experiments is roughly as follows. Conduct an analysis of variance of the results and, if the treatment effects are "significant" sources of variation at some predetermined (arbitrary?) level of probability, a detailed comparison of means is conducted through one of the multiple comparison methods. Once a significant difference between means is ascertained, the economic assessment of this significance is made by budgeting the costs and returns associated with the different treatments and, finally, recommendations are made accordingly.

The chief criticisms of this traditional approach to interpreting experiments are:

(a) it takes no explicit account of previous information such as prior experience or experiments;

(b) it conservatively focuses attention only on making Type I errors, i.e. rejecting a null hypothesis that is true; and

(c) it inevitably depends on probability judgments at arbitrary levels such as ·05, ·01, ·001, etc.

One way of overcoming these difficulties is the use of Statistical Decision Theory or Bayesian Statistics [28]. The essence of the Bayesian approach to interpreting experimental results is that it does explicitly introduce prior knowledge into the analysis. It also takes proper account of both Type I errors and Type II errors (e.g. recommending a new practice that is not really better than the standard) through the use of loss functions which assess the costs of decision errors. Recourse to arbitrary probability judgments is also eliminated. Excellent expositions of Bayesian procedures are available in Schlaifer [30], and Jedamus and Frame [21]. Their application to agricultural research is discussed by Dillon and Officer [13, 27].

It must be noted however, that the development of Bayesian procedures has been controversial and there remains a considerable division of opinion among statisticians as to the acceptability of this statistical philosophy, which *inter alia*, regards the parameters (e.g. the mean) of probability distributions as random variables rather than as fixed but unknown constants. However, general acceptance seems inevitable in the long run and, gradually, Bayesian procedures are being developed for many of the standard tests of conventional statistics (e.g. multiple range tests [15]).

Because the analysis is so much more demanding (of prior information and loss functions), Bayesian analysis of an experiment is a much more complicated business than the traditional method outlined above. Nonetheless it is still the most appropriate method for assessing a decision-oriented experiment, and an analysis which includes any feature of the Bayesian approach (e.g. incorporating prior knowledge) is to be preferred to one that includes none. Hitherto, the agricultural research worker has been advised to feel satisfied that his results were "real" if differences between sample means were "statistically significant and sufficiently large to be of practical or scientific interest" [25, p. 8–9]. The Bayesian approach is to bring these subjective elements of experimental interpretation into the open for ready critical appreciation and thus avoid the pretentions and irrelevant stereotypes of traditional statistical analysis.

3 "HOW MUCH?" EXPERIMENTS

3.1 RESPONSE ANALYSIS

Economists when attempting to influence the design of agricultural research work have stressed the importance of planning experiments so that results are amenable to the marginal analyses of economics [10]. In experiments such as stocking rate trials and fertilizer trials, economic analysis hinges on relating marginal revenue associated with each input to its marginal cost. For an extreme example, when finance is unlimited and uncertainty is absent, an input should be increased until the return to the last unit per acre is just equal to the marginal cost of this unit when the objective is to maximize profits.

One important research problem is to determine the relationship between levels of inputs and marginal revenue. This is generally done by using the concept of a response function which relates level of output, Y, to the levels of variable inputs (see Heady and Dillon [10]). Suppose these are X_1 and X_2 in the two factor case so that the response function is written as $Y = f(X_1, X_2)$, where, for example, the variables might be Y lb wool produced per acre, X_1 number of sheep per acre, and X_2 cwt superphosphate per acre. Many different algebraic forms of response functions have been employed in empirical analyses but by far the most popular have been the second-order polynomial functions of which the "quadratic" function is the simplest:

$$(1) \qquad Y = b_1 + b_2X_1 + b_3X_2 + b_4X_1^2 + b_5X_2^2 + b_6X_1X_2.$$

The empirical popularity of simple polynomials can be explained by their ready estimation through least-squares regression, by the ease with which functions of this class can be algebraically manipulated and, finally, by the widely found adequacy of fit to empirical data. This last point is most important since no response analyst pretends that the complex physical, chemical or biological process on which he is experimenting truly follows a polynomial type of mechanism. It is merely that such functions often suffice to describe the generated data "adequately" over a restricted range of observations of the input variables. For those situations where simple polynomial functions prove inadequate or unduly restrictive, the class of response functions called by Janvry [20] the Generalized Power Production Functions (GPPF) presents useful possibilities which have yet to be exploited in empirical application. This class of functions can be written for several inputs $X_1, X_2, \ldots X_n$ as

$$Y = A \prod_{i=1}^{n} X_i^{f_i(X_1, X_2, \ldots X_n)} e^{g(X_1, X_2, \ldots X_n)},$$

where the $f_i(X_1, X_2, \ldots X_n)$ and $g(X_1, X_2, \ldots X_n)$ denote polynomials of any degree. An example of a two-factor GPPF is given by

$$(2) \qquad Y = b_1 X_1^{b_2 + b_3 X_2} X_2^{b_4} e^{b_5 X_1},$$

In logarithmic form, such equations can also be estimated directly in least-squares regression analysis. The main cost associated with using such relatively flexible response models is the increased algebraic complexity of manipulating the derived equations in economic analyses.

3.2 RESPONSE SURFACE DESIGNS

The economic interpretation of a linear (first order) response function such as $Y = b_1 + b_2X_1$ where $X_1 \geqslant 0$, is that in the absence of an upper constraint on X_1, the optimal quantity to apply is either infinite or zero depending on whether or not it is economic to use the first unit of X_1. As linear relationships are seldom encountered in practice where substantial variations in input levels are involved, such a dilemma is not often met.

Designs for experiments that are intended to yield data suitable for economic analysis should be characterized by sufficient variation in the input levels to provide detectable non-linearity when it is present. This implies that the non-linearity has to be measured by using more than two levels of a factor since it is possible to draw any arbitrary curve through two points. These requirements are most simply met by a design which includes three levels of each factor under investigation. For the case where there is only one factor or variable, the three treatment levels might be replicated in a randomized blocks design.

The complexity of a design increases approximately exponentially with the number of factors involved. In spite of economists' appeals for agricultural experimenters to include many factors (e.g. Dillon and Burley [12]), the most commonly encountered is the two-variable case and only few response experiments involve more than three variables. The simplest adequate design for a two-variable experiment is a 3^2 factorial design. A model such as the "quadratic" function can readily be fitted to data from one replicate of a 3^2 factorial, although a single-replicate experiment would rarely be planned since some replication is necessary to be able to test the estimated relationship for "lack of fit" [14]. In fact, as Mendenhall [24] spells out, a rather more complicated polynomial model including "interaction" terms such as $X_1X_2^2$ can be fitted to this factorial design but in practice it is found that such high-order interactions frequently are of minor statistical importance. They are generally of minor economic importance since the coefficients tend to be small.

This question of the empirical adequacy of second-order polynomial models to describe response phenomena is important because on it hinge the possibilities for using alternative designs to the factorials. In particular, once it is decided that a second-order response function (e.g. equation (1)) will suffice to describe a process over a specified range of experimental variables then several second-order designs [19] are available as possibly lower cost alternatives to the more traditional complete factorial and fractional factorial designs.

The most useful of these designs are the "central composite" and "rotatable" designs. These designs were developed by Box and his associates primarily for use in the chemical industry (Hill and Hunter [19]). Sometimes they are referred to as "response surface" designs but this loose terminology is to be avoided since, as has been noted, a response surface can be fitted to data from factorial designs. Although composite and rotatable designs were originally intended for use in exploring processes characterized by ease and speed of experimental runs and by small error variance, economists such as Heady and Dillon [18] observed their potential usefulness in agricultural research. Economists' main reason for recommending the use of these designs is the economy they provide in reducing the size of multi-factor response experiments. Comparison of replicate sizes for composite designs involving k factors and 3^k factorial and $(1/3)3^k$ fractional factorial designs [18, p. 173] reveals the considerable size reduction that can make response experiments involving five or six factors feasible. Comparable factorial experiments may be intolerably expensive.

Economists' preference for composite designs stems from the viewpoint that most processes worth investigating involve several variable factors and if these are all to be controlled optimally, some account must be taken of the important interactions between the factors. Thus for a given experimental budget, more factors can be studied by using composite designs. Information on interactions can only be established through multi-factor experimentation. However, if only two factors are involved, both composite and 3^2 factorial designs have nine treatments per replicate; so there is little advantage to be gained from using a composite design to economise on experimental resources. There has not been unanimous acceptance of suggestions for adoption of composite and rotatable designs in agricultural research and an example of the debate is provided by the sequence of views presented by Dillon [9], Williams and Baker [32], and Anderson and Dillon [4]. In this exchange, biometricians Williams and Baker took the line that since, for a two-factor experiment, there is little difference in the statistical quality of response functions estimated from either factorial or "response surface" designs, then the factorial designs are to be preferred because they permit a more complete analysis of variance (equivalent to the fitting of a more complex polynomial function). One's posture in this argument depends on the number of response variables being considered and the relative importance one attaches to the complexity of the response model and the completeness of variance decomposition, as opposed to economy in the exploration of a response surface.

Detailed discussion of the construction and use of composite and rotatable designs is provided in most modern books on experimental design including those by Cochrane and Cox [7], Heady and Dillon [18], and Mendenhall [24]. A central composite design consists of a 2^k factorial design augmented by $2k$ "star" points and a "centre" point. It can be made rotatable by a suitable choice of the distance (α) of the "star" points from the centre. For example, if $k = 2$, and $\alpha = 1{\cdot}414$, the design is rotatable and the star points lie on a circle through the factorial points. Other second-order rotatable designs can be constructed by locating treatments at the apices of regular figures of at least $(2^k + k(k - 1)/2)$ sides about a centre point.

The chief advantage claimed for rotatable designs is that the precision of estimation of a response surface follows the same profile independently of the direction moved from the centre. Intuitively this is a desirable feature when interest in the response experiment is equal in all directions from the centre. An appropriate number of replications of the centre point will ensure that the precision (the inverse of variance) is very similar right across the treatment region straddled. However, agricultural economists have seemingly overplayed these desirable features of rotatable designs, since in agricultural experiments there is typically large residual variance and relatively minor differences in the *distribution* of variance over the experimental region are probably unimportant.

To summarize, the following guidelines are suggested for planning response experiments. They have as an underlying tenet the truism of Box [5]—"To find out what happens to a system when you interfere with it you have to interfere with it (not just passively observe it)".

Assess as well as possible from the literature, previous experiments, system models or plain hunch, the number of variable factors (k) which are worthy of inclusion in the experiment.

If $k = 1$, use at least three levels of the factor in a simple experiment perhaps involving qualitative factors such as variety, etc.

If $k = 2$, employ a 3^2 factorial or similar design with the degree of replication determined by consideration of the variability of the process and the precision required in the analysis. There are no simple rules for decision on the replication problem.

If $k \geqslant 3$, and research resources are unlimited, use a 3^k factorial or a fraction of this. If there are constraints on the resources available to the experimenter such as cash, experimental materials, suitable land, etc., consider the use of a central composite design and, less importantly, consider making it rotatable.

No matter which design is chosen, the experimenter is faced with the problem of deciding where to centre the design. Generally, the centre is best located at what prior reasoning indicates is the most likely "most economical" operating point. Such a point is seldom if ever obvious and usually boils down to a good guess at what is at any rate an ephemeral and a personal optimum. Again, determining the treatment region to straddle with an experiment is difficult and must be guided by consideration of the biological and economic constraints on the ranges of variables along with the knowledge that the wider is the straddle, the less likely will be the descriptive adequacy of second-order polynomial and other simple response models.

3.3 INTERPRETATION OF REGRESSION STATISTICS

Brief mention needs to be made of some of the problems of interpreting regression statistics because there is an apparent inconsistency between what has been recommended above for interpreting "Yes/No" experiments and the present practice of applied regression analysis. Unfortunately, the development of Bayesian methods for regression problems has not kept pace with the developments for the intrinsically simpler problems such as comparing means. The techniques of Bayesian regression analysis so far available are very artificial in the sense of requiring unrealistic but convenient assumptions about the prior and sample distribution functions. The use of Bayesian methods in interpreting "How much?" experiments will probably not be feasible at least until the late 1970's, but progress is encouraging [34].

Meantime there are some "Bayesian-type" suggestions that can be made for interpreting regression equations. The usual practice is to fit as simple a response model as possible and, on the basis of an "F" test of the overall regression and the "t" tests on the "significance" of the results. Typically a coefficient or an effect is reported as being "significant at the ·05 (or ·01 etc.) level". Of course, such statements are devoid of any economic content, but more importantly they are sometimes ill-founded on irrelevant hypotheses. Hypotheses under test in regression analysis are seldom stated explicitly but, for example, the usual implied null hypothesis in a "t" test is that the "population coefficient" is zero and the alternate hypothesis is that this coefficient is not zero. Analysts should give careful consideration to hypothesis formulation and thus be cautious in accepting results from computer programs which churn out answers only to the usual test. Occasionally, a more valid null hypothesis (based on prior knowledge) will be that the population coefficient has some particular non-zero value, in which case the "t" test can be simply modified.

A common procedure adopted by regression analysts is to fit a supposedly complete response model to the experimental data and then simply refit a reduced model which excludes those variables whose coefficients did not reach some arbitrary level of "significance". This is not good practice but may be condoned when the analyst is fully aware of the implications of the reduced model. Suppose in fitting a model such as the two-factor quadratic the terms for X_1^2 and X_1X_2 are dropped because they were "not significant at the ·05 level". The analyst is then saying that response to X_1 is linear (with all that that implies) and that there is zero interaction between X_1 and X_2. These are strong conclusions to reach from an economic point of view and the cost of possible Type II errors may be substantial. The sacred ·05 probability of a Type I error is seldom good science or good economics.

A regression technique which has been used increasingly is automatic computer selection of the "best" regression equation. Some of the several methods in use are described by Draper and Smith [14], but all should be avoided by agricultural response analysts since they implicitly permit the sample data to generate their own hypotheses which are then examined in completely arbitrary and mechanical fashion.

4 REPORTING EXPERIMENTS

Only very broad generalizations can be made about experimental reporting because of the diversity of experiments conducted and the range of economic interpretations possible. Some research is fairly adequately reported in the usual scientific format. On the other hand, the usual style of reporting is not always adequate to provide the economist with all the information he requires for, say, an economic assessment of the results. Three main areas where more attention is required can be identified.

4.1 MEASUREMENT AND REPORTING OF EXOGENOUS VARIABLES

It has always been regarded as good scientific reporting to describe the experimental environment in fairly complete detail. For example, variety of plant or animal, soil type, soil fertility, measurements and measurements of weather variables for the duration of an experiment, etc. are commonly reported, but it is surprising how frequently an important measurement is absent, such as mortality, supplementary feeding details, forage nitrogen determinations, etc. It is also important that an economist charged with analysis of an experiment within a farm management context be given full information about the circumstances of a trial. If there is uncertainty about the appropriate amount of detail to report on exogenous variables it is better to err on the side of too much.

Measurement and reporting of exogenous variables allows for their subsequent inclusion in more comprehensive models of a system. For instance, a series of experiments might lead to the progressive enlargement of an algebraic response model, but this pooling of individual response functions would only be potentially possible if all the relevant variables were measured for all experiments. This evolution of a larger model can be through a continuous algebraic model such as described by Byerlee and Anderson [6] or may encompass only a few discrete states such as several seasonal categories. Either way, the techniques of decision theory reviewed by Dillon [11] can be employed to trace the evolution of the research systematically, to assess the value of continued experimentation, and to indicate farmers' best bet decisions based on the experimental information.

4.2 REPORTING SUFFICIENT STATISTICS

Sufficient statistics are those that summarise all the information from a sample concerning the parameters so that any additional statistics are uninformative. The concept was introduced by R. A. Fisher and is important in both classical and Bayesian interpretation of experiments. It is essential that sufficient statistics be reported for all users of research results—be they economists or other scientists. Bayesian users are especially hindered by bad reporting since they require the sufficient sample statistics to combine with their prior information to arrive at a posterior distribution for the investigated process. It is unforgivable to report results in such a way that the sufficient statistics are not recoverable. Even when recovery is possible, the user is not well served if he is forced to calculate these numbers which might have been given him directly (see Raiffa and Schlaifer [29]). Too often (particularly in the *Australian Veterinary Journal*) the reader is presented with only a table of means and a variety of symbols denoting the levels of "statistical significance".

What constitutes sufficient statistics in any case depends on the process, the statistical technique, and the nature of the probability distributions assumed to be operative. For the majority of designed "Yes/No" experiments, sample sizes, means and variances are sufficient. For the "How much?" experiments under the conventional assumptions of normally distributed disturbances, it is sufficient to report the (vector of) estimated regression coefficients, the associated variance-covariance matrix, the error mean square and the sample size [2, p. 183]. Unfortunately, reporters often merely give the regression coefficients adorned with symbols denoting "significance". A procedure that is just tolerable in that sufficient statistics can be calculated when required is to report sample size, the coefficient of multiple determination R^2, the regression coefficients and their standard errors, the standard error of the estimate, and either the zero-order correlation matrix of regressors or precise details on the experimental design matrix. Whenever there is a doubt as to the sufficiency of reporting, a complete tabulation of the data should be given in an appendix.

4.3 MEASUREMENT OF RISK IN EXPERIMENTS

Agricultural economists have over recent years become increasingly aware of the importance of risk in decision making on farms (e.g. [11, 26]). The fact that farmers as individuals are usually not indifferent to risk goes some way towards explaining their apparent aberrant behaviour in not adopting practices which on technical and expected profit grounds seem worthwhile. It may be that in many cases the new practices are *not* more risky than standard practices; but farm decision makers can only make decisions on the basis of their subjective appreciations of risk and in the light of their unique attitudes to risk.

The implications for reporting agricultural research are clear. As far as possible, measures of risk observed in experimental data should be explicitly reported [1]. One of the problems here is that economists have not agreed on what constitutes an acceptable measure of risk. The ideal report would completely specify the probability distribution for the running of the process under experimentation [22], but this ideal would seldom be feasible. More pragmatic assessments of risk have to be made. Variance, say of process profit, has some unsatisfactory features as an economic measure of risk but is one of the more easily estimated measures. Even so, it is generally appreciated that estimation of variances from small samples is much less robust (i.e. more sensitive to sampling errors and underlying population distributions) than is estimation of means. An example of the use of variance induced by the operating level of variable factors of production is provided by the work of McArthur and Dillon [23] on optimal sheep stocking rates.

While it is recognized that risk cannot yet, and probably never will, be susceptible to unambiguous assessment, the best approach is that some description of the risk encountered in experimental work is better than none. For example, results for individual years and seasons should be reported as well as long-run averages.

5 OPTIMAL STOPPING

One of the desirable features of the Bayesian approach to statistical decision making is that it does provide an answer to the question of when to stop experimenting. In Bayesian jargon, experimental work should cease when the marginal cost of further experimentation exceeds the expected value of sample information [11, 30]. However, as has been observed, there are serious practical difficulties associated with employing the Bayesian approach in many experimental situations in agriculture so that even for many "Yes/No" experiments it is seldom clear when a research programme should be wound up.

The situation for deciding when to terminate "How much?" experiments is not so hopeful, although a start has been made on answering this question by Anderson and Dillon [3] and some further suggestions have been made by Seagraves [31]. Economists have for years been exhorting agricultural research workers to conduct response-surface rather than "point-estimate" experiments and it behoves them to give more guidance on the likely payoffs from such experiments, and on when they should cease.

One possibility for insights to optimal stopping is to channel research results into simulation models of systems that are under experimental investigation. Simulation models of the type used in agricultural systems research [8] can provide a unifying view ot research progress, an assessment of when a gap has been adequately filled and an indication of components of the system that are still insufficiently understood. However, although it seems that incorporation of the systems approach in the conduct and evaluation of experiments could lead to improvement and rationalisation in research, past experience counsels that there can be no simple panacea to the problems besetting agricultural research.

6 SUMMARY

The main points made above can be briefly summarized as follows:

(i) Interpretation of experiments involving comparison of means can be improved and made more relevant for extension purposes by incorporating elements of Bayesian analysis (namely, prior information, subjective probabilities, costs of errors of the first and second kind).

(ii) Experimenters exploring response to three or more variables should make more use of composite designs.

(iii) Regression statistics should be interpreted subjectively and in terms of economic implications and previous experience rather than in arbitrary mechanical fashion.

(iv) Experimental reporting can often be improved with respect to more detail on non-treatment variables, sufficient statistics, assessment of intrinsic risk, and risk induced by treatments.

(v) The problem of optimal stopping of experimental programmes is an important aspect of research deserving more research.

REFERENCES

[1] ANDERSON, J. R., "Economic Interpretation of Fertilizer Response Data", this *Review*, Volume 35, No. 1 (1967), pp. 43–57.

[2] ANDERSON, T. W., *An Introduction to Multivariate Statistical Analysis*, (New York: Wiley, 1958).

[3] ANDERSON, J. R. and DILLON, J. L., "Economic Considerations in Response Research", *American Journal of Agricultural Economics*, Volume 50, No. 1 (February, 1968), pp. 130–142.

[4] ANDERSON, J. R. and DILLON, J. L., "A Comparison of Response Surface and Factorial Designs in Agricultural Research: Comment", this *Review*, Volume 37, No. 2 (June, 1969), pp. 130–132.

[5] BOX, G. E. P., "Use and Abuse of Regression", *Technometrics*, Volume 8, No. 4 (November, 1966), pp. 625–629.

[6] BYERLEE, D. R. and ANDERSON, J. R., "Value of Predictors of Uncontrolled Factors in Response Functions", *Australian Journal of Agricultural Economics*, Volume 13, No. 2 (December, 1969), pp. 28–37.

[7] COCHRANE, W. G. and COX, G. M., *Experimental Designs*, 2nd Edn. (New York: Wiley, 1957).

[8] DENT, J. B. and ANDERSON, J. R., Eds, *Systems Analysis in Agricultural Management*, (Sydney: Wiley, 1971).

[9] DILLON, J. L., "Economic Considerations in the Design and Analysis of Agricultural Experiments", this *Review*, Volume 34, No. 2, (June, 1966), pp. 64–75.

[10] DILLON, J. L., *The Analysis of Response in Crop and Livestock Production*, (Oxford: Pergamon, 1968).

[11] DILLON, J. L., "An Expository Review of Bernoullian Decision Theory in Agriculture: Is Utility Futility", this *Review*, Volume 39, No. 1 (March, 1971), pp. 3–80.

[12] DILLON, J. L. and BURLEY, H. T., "A Note on the Economics of Grazing and its Experimental Investigation", *Australian Journal of Agricultural Economics*, Volume 5, No. 2 (December, 1961), pp. 123–32.

[13] DILLON, J. L. and OFFICER, R. R., "Economic v. Statistical Significance in Agricultural Research and Extension", *Farm Economist*, (1971), Volume 12, No. 1 (1971), pp. 1–15.

[14] DRAPER, N. R. and SMITH, H., *Applied Regression Analysis*, (New York: Wiley, 1966).

[15] DUNCAN, D. R., "A Bayesian Approach to Multiple Comparisons", *Technometrics*, Volume 6, No. 2 (May, 1965), pp. 171–222.

[16] HEADY, E. O., *Economics of Agricultural Production and Resource Use*, (Englewood Cliffs: Prentice-Hall, 1952).

[17] HEADY, E. O., *Agricultural Problems and Policies of Developed Countries*, (Oslo: Bondenes Forlag, 1966).

[18] HEADY, E. O. and DILLON, J. L., *Agricultural Production Functions*, (Ames: Iowa State University Press, 1961).

[19] HILL, W. J. and HUNTER, W. G., "A Review of Response Surface Methodology: A Literature Survey", *Technometrics*, Volume 8, No. 4, (November, 1966), pp. 571–590.

[20] JANVRY, A. DE, "The Class of Generalized Power Production Functions", unpublished manuscript, Projecto Pro Economica Agraria, Buenos Aires, 1971.

[21] JEDAMUS, P. and FRAME, R., *Business Decision Theory*, (New York: McGraw-Hill, 1969).

[22] LLOYD, A. G., "Agricultural Experiments and their Economic Significance", this *Review*, Volume 26, No. 3 (September, 1958), pp. 185–209.

[23] MCARTHUR, I. D. and DILLON, J. L., "Risk, Utility and Stocking Rate" *Australian Journal of Agricultural Economics*, Volume 15, No. 1 (April, 1971), pp. 20–35.

[24] MENDENHALL, W., *Introduction to Linear Models and the Design and Analysis of Experiments*, (Belmont. Calif: Wadsworth, 1968).

[25] MOULE, G. R., Ed., *Field Investigations with Sheep: A Manual of Techniques*, (Melbourne: C.S.I.R.O., 1965).

[26] OFFICER, R. R. and ANDERSON, J. R., "Risk, Uncertainty and Farm Management Decisions", this *Review*, Volume 36, No. 1 (March, 1968), pp. 3–19.

[27] OFFICER, R. R. and DILLON, J. L., "Probability and Statistics in Agricultural Research and Extension", *Journal of the Australian Institute of Agricultural Science*, Volume 34, No. 3 (September, 1968), pp. 121–129.

[28] RAIFFA, H., *Decision Analysis*, (Reading: Addison-Wesley, 1968).

[29] RAIFFA, H. and SCHLAIFER, R., *Applied Statistical Decision Theory*, (Boston: Harvard Business School, 1961).

[30] SCHLAIFER, R., *Probability and Statistics for Business Decisions*, (New York: McGraw-Hill, 1959).

[31] SEAGRAVES, J. A., "Economic Considerations in Response Research: Comment", *American Journal of Agricultural Economics*, Volume 52, No. 4 (November, 1970), pp. 607–609.

[32] WILLIAMS, R. J. and BAKER, J. R., "A Comparison of Response Surface and Factorial Designs in Agricultural Research", this *Review*, Volume 37, No. 2 (June, 1968), pp. 165–177.

[33] WRAGG, S. R., "Co-operative Research in Agriculture and the Provision of Input/Output Coefficients", *Journal of Agricultural Economics*, Volume 21, No. 1, (January, 1970), pp. 85–98.

[34] ZELLNER, A., *Introduction to Bayesian Inference in Econometrics* (New York: Wiley, 1971).