34th Annual Conference of the Australian Agricultural Economics Society,

University of Queensland, Brisbane, 12-15 February 1990

The Decision to Fund a Research Proposal: Some Results for Wool Production

A. C. Wohlers,
La Trobe University

Chris Vlastuin,
Australian Wool Corporation

Abstract

In this paper the determinants of the decision to fund a research proposal are examined using Production Research Advisory Committee (PRAC) data. The committee is a subcommittee of the Wool Research and Development Council (WRDC). A variety of descriptive (correlation, principal component and factor analysis) and inferential (standard multiple linear, truncated, logistic regressions and discriminant analysis) techniques are applied to proposal data and the associated PRAC evaluations. These techniques are applied in order to examine the extent to which research proposal data can be quantified, simplified and PRAC decisions predicted. The report concludes that PRAC decisions are determined to some extent by two underlying characteristics of the proposal data: 'demand' information which relates to industry needs and 'supply' information which relates to available academic/institutional skills in the various research areas. These findings are the basis for changes to the referee reports sent out to reviewers.

1 Introduction

One of the areas of interest to the Research and Development Department of the Australian Wool Corporation (AWC) is the application of analytical techniques as tools for decision making in the AWC with a view to increasing efficiency and cutting costs.

In this case, it was decided to examine the process of selection for funding of research proposals submitted to the Wool Research and Development Council (WRDC) and evaluated by the WRDC's Production Research Advisory Committee (PRAC). Between 200 to 250 new proposals are considered for funding each year during the annual budget meeting of PRAC. This committee is one of five committees responsible for wool-related research and is funded by the Wool Research and Development Fund (WRDF) which is financed by wool growers and the Australian Government. About $23 million of the total budget of $56 million will be spent on production research in 1989/90. Each application evaluated by PRAC contains both budgetary and technical information. In addition, the Committee has at least one report on each proposal from external referees. Referees are asked to rate each proposal according to how it meets certain criteria and are able to provide additional comments if they desire. Consequently, the committee faces a substantial volume of material which must be processed in a short time, so any techniques which:

- distinguish determining factors in the decision to fund a proposal,
- use this information to rank proposals so as to distinguish between particularly 'good' and particularly 'bad' proposals, and borderline cases which may require further examination, and
- forecast future funding decisions,

may ease the work load. Statistical techniques were chosen because some of the information used by PRAC can be easily quantified.

This quantification comes about because PRAC members score each proposal between 1 and 4 during the annual budget meeting and given that there are 12 members, each proposal will score between 12 and 48. A low score indicates a preferred proposal. In 1989/90, proposals with a PRAC score of less than or equal to 33 were recommended for funding. A referee's report (see Appendix 1) consists of 15 questions, all of which use a categorical rating scheme (mostly five levels). In addition, proposal cost and source (i.e. CSIRO, tertiary institution etc.) information can be used. Consequently, models can be constructed by considering the PRAC score or the decision to fund or not to fund as being related to the referees' question scores and the cost and source information. More formally the PRAC score or the implicit funding decision is an observation on a *dependent* variable and the other information forms a set of observations on a number of *explanatory* variables. The statistical models in this report are different ways of describing the interaction of these variables that enable the investigation of the following issues.

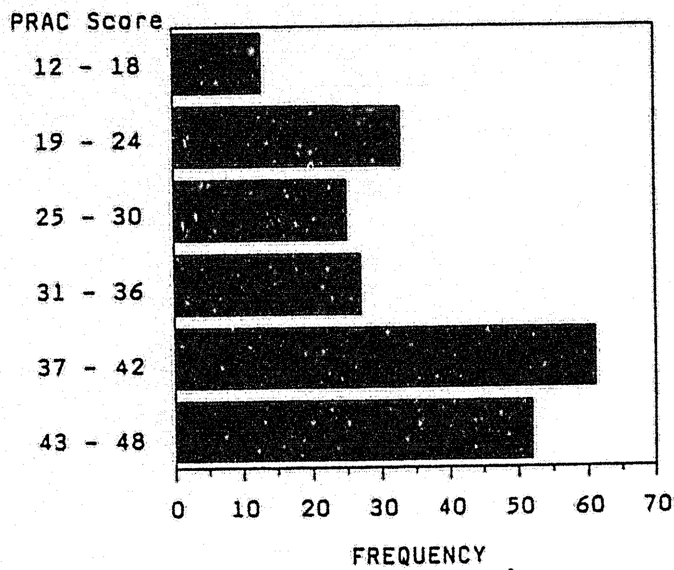
- Is any of the quantifiable information used by the committee to arrive at its decisions ?
- If it is, which information is most influential ?
- Which decisions are unusual or at variance with other decisions ?
- Can proposals be ranked prior to a PRAC meeting ?
- Or can the data be summarised or aggregated in a way that does not mask valuable information?

2 The Data

The sample data consists of 472 sets of observations on 220 new proposals considered for funding in 1989/90. Scores on individual questions were averaged in instances when more than one referee's report was available per proposal. Nine proposals were deleted from the sample because referee reports were partially completed or because the PRAC score on the proposal was unavailable. This left 211 observations for analysis.

Data for variables coming from the referee questionnaire (see Appendix 1) required some processing prior to use because of the nature of the individual questions. If the answer to Question 10 was *no* then a score of 6 was assigned. Question 14 provides an overall project ranking. The A=excellent, B=very good, C=good and D=fair rankings were converted into numerical scores from 4 to 1, respectively. Question 15 was transformed using the reciprocal of the project ranking score. The higher this score the more favourable the project would appear to be. Some variables were suitable for the creation of dummy variables like goal number or project number. For instance the prefix CS on the project number designates that the proposal comes from CSIRO. Hence the variable CS takes two values: 0=non-CSIRO proposal, 1=CSIRO. Such a variable can be used to assess whether some research institutions are favoured over others when considering proposals. Data on the cost of each proposal is also available.

A starting point for the analysis is to examine the distribution of the PRAC scores in the sample. As these scores constitute the dependent variable, a knowledge of their distribution might have some bearing on how the models may be expected to perform. The main point to note about



FREQ	CUM. FREQ	PERCENT	CUM. PERCENT
13	13	6.16	6.16
33	46	15.64	21.80
25	71	11.85	33.65
27	98	12.80	46.45
61	159	28.91	75.36
52	211	24.64	100.00

Figure 1: Histogram of PRAC Scores

Figure 1 is the number of proposals near the cut off point i.e., 13 percent or a total of 27 projects fall within 3 points of 33. Obviously the various models are likely to do poorly in trying to predict the committee's decision for these 'marginal' projects. Conversely it is hoped that these models will do comparatively well in predicting the Committee's decision on say the top/bottom 30 percent of proposals.

3 Exploratory Data Analysis (EDA)

3.1 Examination of Sample Correlation Matrix

Although Figure 1 gives some idea of how the sample of PRAC score values are distributed it gives no idea how this distribution is related to the supposed explanatory variables. It would be possible to look at the distribution of each of the explanatory variables in a similar fashion or look at scatterplots of pairs of variables, however given the number of variables in the PRAC data this would be rather tedious and would not reveal how more than any given pair of variables interact. One technique that does provide an insight into how all the variables interact with one another is the examination of the *sample correlation matrix* which is not only valuable on its own but is very important in many of the techniques described below.

VARIABLE	PRAC SCORE	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	CS	COST
PRAC_SCORE		-41	-46	-40	-52	-48	-50	-51	-42	5	-33	-38	-29	-49	-61	-43	-14	16
Q1	-41		46	45	39	43	45	52	62	7	18	53	53	43	52	37	-12	-15
Q2	-46	46		59	62	64	52	49	45	20	43	27	22	51	71	42	13	-1
Q3	-40	45	59		64	64	57	42	33	11	33	36	25	51	61	35	6	-13
Q4	-52	39	62	64		69	58	39	37	14	42	28	30	53	63	35	6	-8
Q5	-48	43	64	64	69		59	51	39	15	39	38	28	54	70	44	11	-9
Q6	-50	45	52	57	58	59		50	32	15	23	46	38	52	62	28	3	-14
Q7	-51	52	49	42	39	51	50		55	5	21	43	38	45	57	35	-8	-38
Q8	-42	62	45	33	37	39	32	55		16	21	39	49	41	57	35	-11	-9
Q9	5	7	20	11	14	15	15	5	16		8	-5	1	1	14	1	-5	17
Q10	-33	18	43	33	42	39	23	21	21	8		12	-2	28	40	26	7	-4
Q11	-38	53	27	36	28	38	46	43	39	-5	12		57	30	42	22		-20
Q12	-29	53	22	25	30	28	38	38	49	1	-2	57		46	38	24	-20	-27
Q13	-49	43	51	51	53	54	52	45	41	1	28	30	46		63	40	3	-14
Q14	-61	52	71	61	63	70	62	57	57	14	40	42	38	63		62	14	-12
Q15	-43	37	42	35	35	44	28	35	35	1	26	22	24	40	62		10	-7
CS	-14	-12	13	6	6	11	3	-8	-11	-5	7		-20	3	14	10		23
COST	16	-15	-1	-13	-8	-9	-14	-38	-9	17	-4	-20	-27	-14	-12	-7	23	

In order to increase readability of the sample correlation matrix, in the above table leading zeros and decimal points have been omitted, the 1's in the diagonal have been suppressed and only two decimal places are shown hence -41=-0.41, non-diagonal blank implies $r_{ij} < 0.01$. The interpretation used here follows the examples found in Section 3.2.1 of Chatfield and Collins (1980). The sample correlation matrix for the PRAC data is best examined by considering the squared correlations. The r_{ij}^2 is an estimate of the proportion of the variance of X_i explained by X_j. Consequently if $r_{ij} \leq 0.7$ then the correlation will account for less than 50% of the observed variation in X_i. The level of correlation considered of *practical importance* is up to the investigator. Chatfield and Collins (1980) suggest values above 0.7 may be considered large. In this section and in subsequent sections where sample correlations are discussed, only $r_{ij} \geq 0.5$ will be considered. Smaller values imply that the correlation accounts for less than 25% of the observed variation.

One general comment about the PRAC data sample correlations is that there are none greater than 0.71. Furthermore the largest correlations are between explanatory variables (Referee question scores: Q14 and Q2, Q14 and Q5, Q4 and Q5) rather than between the dependent variable (PRAC score) and the explanatory variables. This is further evidence that the PRAC data is fairly 'noisy' and any models built from this data may have trouble distinguishing all but the best (or the worst) proposals. Q9 has the lowest correlation with the PRAC score and its cross-correlation with other questions in the referee report is also comparatively low. Its correlations with the PRAC score, like the project cost variable, is incorrectly signed however given their small absolute value they are probably not significantly different from zero. Interestingly, Q14 the overall project rating given by the referee is most highly correlated with Q2 and Q5. These questions relate to scientific merit of the proposal and the method of analysis, respectively. This result perhaps is not surprising given that research scientists are asked to assess the work of their colleagues.

The CS and Cost variables have low correlations with each other and all other variables. As a result of these observations on the sample correlation matrix the rest of this section will examine the structure between the referee questions (variables Q1 to Q15). These variables represent the bulk of the data and appear to contain the highest correlations. These correlations are suggestive of relationships between the referee questions which require further analysis. Such analysis may in turn lead to some simplification or summary of the information contained in the referee questions. Consequently the following EDA techniques emphasize *dimensional reduction* (i.e. re-expressing the data with a smaller set of variables) and its interpretation.

3.2 Principal Component Analysis (PCA)

The following description is taken from Afifi and Clark (1984) pp309-310:

> Principal components analysis is performed in order to simplify the description of a set of variables....The technique can be summarized as a method of transforming the original variables into new, uncorrelated variables. The new variables are called the *principal components*. Each principal component is a linear combination of the original variables. One measure of the amount of information conveyed by each principal component is its variance. For this reason the principal components are arranged in order of decreasing variance. Thus the most informative principal component is the first and the least informative is the last.....
>
> An investigator may wish to reduce the dimensionality of the problem, i.e., reduce the number of variables without losing much of the information. This objective can be achieved by choosing to analyse only the first few principal components. The principal components not analyzed convey only a small amount of information since their variance is small. This technique is attractive for another reason, namely, that the principal components are not intercorrelated. Thus instead of analysing a large number of original variables with complex interrelationships, the investigator can analyse a small number of uncorrelated principal components.

3.2.1 Theory

Consider a set of variables $X_1, X_2, \ldots, X_k$, PCA produces a new set of variables $Y_1, Y_2, \ldots, Y_k$ (*principal components*) which are linear combinations of the original variables;

$$\left.\begin{array}{rcl} Y_1 & = & a_{11}X_1 + a_{12}X_2 + \ldots + a_{1k}X_k \\ Y_2 & = & a_{21}X_1 + a_{22}X_2 + \ldots + a_{2k}X_k \\ \vdots & = & \vdots \\ Y_k & = & a_{k1}X_1 + a_{k2}X_2 + \ldots + a_{kk}X_k \end{array}\right\} \qquad (1)$$

The coefficients a_{ij} are chosen so that;

- $\text{Cov}(Y_i, Y_j) = 0$, for $i, j = 1, \ldots, k$ and $i \neq j$; the principal components are *uncorrelated*,
- $\text{Var}(Y_1) \geq \text{Var}(Y_2) \geq \ldots \geq \text{Var}(Y_k)$, they are in order of decreasing variance,
- $\sum_{i=1}^{k} \text{Var}(Y_i) = \sum_{i=1}^{k} \sigma_{ii}^2$; the sum of the variances of the principal components equals the variances of the original variables.

The total variation in the data (as measured by its variance) cannot be easily partitioned amongst the original variables because they are almost certainly correlated. However the principal components are uncorrelated so that their variances (called *eigenvalues* extracted from the correlation matrix) partition the observed variation in the data. By ranking the Y_i's by their variances their relative importance to the total variance of the data can be assessed. The original variables could also be ranked in this way but because they are correlated the relative size of their contribution to the observed variation is hard to unravel. It is important to realize that principal components of a set of variables depend critically upon the scales used to measure the variables. This is not too surprising when one considers how the principal components are defined, as a linear combination of variables. Clearly the loadings would change if the units of the variables were changed so that they had a different numerical relationship to one another. As a consequence of this PCA is usually based on standardised variables (to standardise a variable involves dividing it by its standard deviation) which all have the same scale. Standardisation has the effect of making PCA treat each variable on the same footing.

3.2.2 Results of PCA on PRAC data

The application of PCA to the PRAC data involved finding principal components amongst the variables relating to the referee questions (Q1 to Q15). The aim was to see if a few principal components could summarise the information in these variables and whether these components could be usefully interpreted in the context of possible criteria of proposal success. An analysis of the PRAC data revealed that almost 60 per cent of the observed variance is accounted for by the first two components. The other components individually account for very little of the variance. The problem is deciding which to discard. Clearly the first two components are important. But how many more components should be retained? A useful graphical tool in this context is a *scree plot* (see Figure 2). In this method, named after the rubble at the bottom of a cliff, the eigenvalues of each component are plotted in order of size, and then the elbow in the curve is identified by applying, say, a straightedge to the bottom proportion of the eigenvalues to see where they form a straight line. The number of components retained is given by the point at which the components curve above the straight line formed by the smaller eigenvalues (Dillon and Goldstein (1984) p48-

49). In Figure 2 the method clearly suggests retention of the first two components

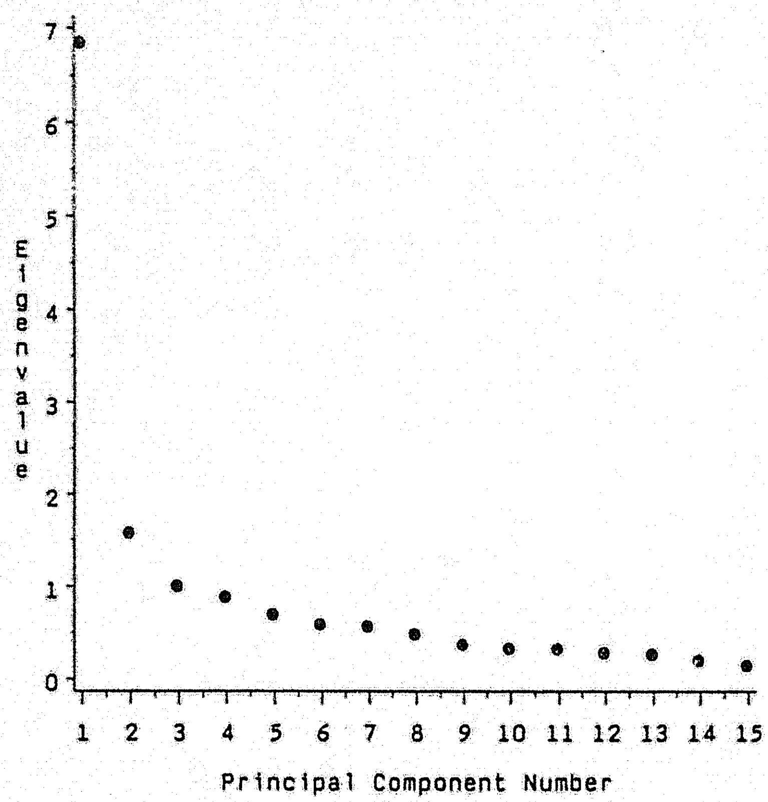


Figure 2: Plot of Eigenvalues versus Principal Component Number

Variables X_i	Description (See Appendix I)	Y_1 (a_{i1})	Y_2 (a_{i2})
Q1	'.. is directly applicable to identified .. R&D goal'	0.270	-.307
Q2	'.. significant scientific merit'	0.300	0.241
Q3	'.. objectives clearly stated'	0.284	0.177
Q4	'.. understanding of topic indicated'	0.290	0.244
Q5	'.. methods .. appropriate'	0.308	0.197
Q6	'.. personnel/facilities adequate for project'	0.283	0.015
Q7	'.. costs reasonable in relation to .. benefits'	0.269	-.162
Q8	'.. increases in productivity to woolgrowers..'	0.254	-.250
Q9	'.. results .. applicable to entire .. industry'	0.063	0.216
Q10	'.. duplication ? .. warranted ?'	0.170	0.395
Q11	'Communication to .. users .. easily achieved'	0.220	-.405
Q12	'.. short time before findings .. adopted..'	0.209	-.499
Q13	'.. probability of achieving project objectives..'	0.277	-.002
Q14	Overall Project Rating	0.335	0.093
Q15	Project Ranking	0.222	0.065

For the sake of brevity and ease of interpretation the approach taken is to examine only the first two eigenvectors (Y_1 and Y_2 which account for 45.9% and 10.7% of the variance respectively). With the PRAC data set it was felt that having to analyse more than two principal components given the exploratory nature of the investigation was unjustified. The effort required to interpret a solution of more than two dimensions was not considered worth the cost. Y_1 has relatively high loadings for variables Q1 to Q8, Q13, and Q14. Q1 to Q6, and Q13 are questions that could be construed as related to the scientific or academic quality of the proposal. Variable Q14 is an overall proposal ranking. Vlastuin (1989) has described these questions as measuring the quality of the *supply* of research from scientists and relevant research institutions. Variables Q8 to Q12 can be taken as measuring usefulness of a proposal to the farmers and the wool growing industry as a whole (i.e. ease of communication, range and speed of application etc.). Vlastuin (1989) has described these questions as measuring the *demand* for research. The demand questions as a group do not load as heavily on Y_1 as they do on Y_2 (especially Q11 and Q12). Another point to note about Y_2 is that the supply variable loadings have positive signs whereas the demand variable loadings tend to have negative signs. Consequently, Y_1 can be seen to describe the supply attributes of a proposal and Y_2 *contrasts* the supply and demand attributes of a proposal. To examine Y_1 and Y_2 in more detail their scores for each observation has been plotted in Figure 3. In addition to displaying the

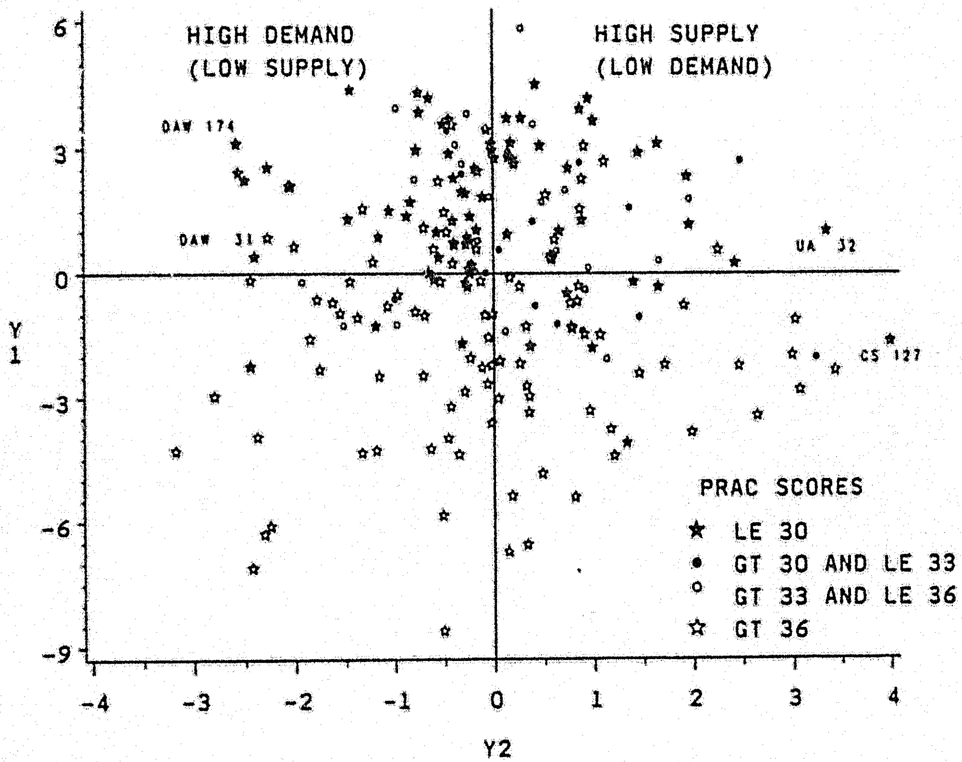


Figure 3: First Principal Component (Y1) versus the Second Principal Component (Y2)

first two principal component scores this figure uses a symbol code that indicates membership of a range of PRAC scores. Closed stars represent the 'very good' projects which receive a score of 30 or less. Open stars represent the 'very bad' projects that receive a score greater than 36. Marginal projects are denoted by circles, closed circles are successful (score from 31 to 33) and open circles unsuccessful (score from 34 to 36). It is apparent from the Figure that considerable overlap still exists across the four groups, nevertheless some structure is evident with the 'very bad' projects figuring predominently in the bottom half of the graph and the 'very good' projects in the upper half, some outliers are also evident. An interesting result is obtained if one examines the scatter of the points as they relate to the horizontal axis. Note that Y_2 has a range from approximately -2.5 to 3.5 suggesting that a decision to fund a proposal is influenced more by 'supply' than 'demand' type characteristics. More interesting is the fact that the marginal but funded proposals (closed circles) lie predominently on the right hand side of the vertical reference line suggesting that supply factors dominate. Further analysis along these lines with particular reference to individual outliers is found in Vlastuin (1989).

Projects lying at the extreme ends of the Y_2 axis have been identified to further illustrate the idea of 'demand' versus 'supply' characteristics of a proposal and their associated scores for each question are given in the following table. Mean scores for each question for the 211 projects are also reported. 'S' and 'D' identify questions which are of primary interest to the researcher and the woolgrower, respectively and relate back to the sign of the coefficients for Y_2 given above in the table of the first two principal components.

Variable X_i	Description (See Appendix I)	Mean Score	Attribute	Supply Driven		Demand Driven	
				CS 127	UA 32	DAW 174	DAW 31
Q1	'.. is directly applicable to identified .. R&D goal'	4.2	D	2.6	3.3	5.0	5.0
Q2	'.. significant scientific merit'	3.7	S	4.0	4.7	4.0	3.0
Q3	'.. objectives clearly stated'	3.9	S	4.0	4.7	4.0	3.5
Q4	'.. understanding of topic .. indicated'	3.7	S	4.3	4.3	5.0	2.5
Q5	'.. methods .. appropriate'	3.6	S	3.3	4.7	5.0	3.0
Q6	'.. personnel/facilities adequate for project'	4.0		3.7	5.0	5.0	4.0
Q7	'.. costs reasonable in relation to .. benefits'	3.6	D	3.0	3.7	5.0	5.0
Q8	'.. increases in productivity to woolgrowers..'	3.4	D	2.3	2.3	3.0	4.0
Q9	'.. results .. applicable to entire .. industry'	3.0	S	4.3	2.0	1.0	1.5
Q10	'.. duplication?.. warranted?'	5.1	S	6.0	6.0	4.0	6.0
Q11	'Communication to .. users .. easily achieved'	4.7	D	1.7	2.7	5.0	5.0
Q12	'.. short time before findings .. adopted..'	2.8	D	1.3	1.3	5.0	5.0
Q13	'.. probability of achieving project objectives..'	0.6		0.7	0.8	1.0	0.7
Q14	Overall Project Rating	2.4		2.3	3.3	2.0	2.5
Q15	Project Ranking	2.3		2.3	1.6	2.0	2.1

Examination of this table clearly demonstrates that projects 'CS 127' and 'UA 32' while very good scientific research proposals are less likely to lead to significant spin-offs to woolgrowers. Conversely, projects 'DAW 174' and 'DAW 31' are less research oriented, but directly address problems faced by woolgrowers.

3.3 Factor Analysis (FA)

FA derives new variables called *factors* which try to explain the interrelationships amongst variables. It may seem very similar to PCA however it approaches the issue of problem simplifcation from the other 'direction'. PCA seeks to simplify the problem by reducing the dimensionality of the *data* and reexpressing it in terms of uncorrelated variables (principal components). FA tries to simplify the problem by expressing the *variables* by a smaller set of factors which are seen as generating the observed variables.

3.3.1 Theory

Consider a set of variables $X_1, X_2, \ldots, X_k$ which have been standardised. The object of FA is to represent each of these variables as a linear combination of a smaller set of *common factors* $(F_1, F_2, \ldots, F_m)$ plus a factor unique to each variable $(e_1, e_2, \ldots, e_k)$:

$$\left.\begin{array}{rcl} X_1 & = & l_{11}F_1 + l_{12}F_2 + \ldots + l_{1m}F_m + e_1 \\ X_2 & = & l_{21}F_1 + l_{22}F_2 + \ldots + l_{2m}F_m + e_2 \\ \vdots & = & \vdots \\ X_k & = & l_{k1}F_1 + l_{k2}F_2 + \ldots + l_{km}F_m + e_k \end{array}\right\} \quad (2)$$

where the following assumptions are made:

- m is the number of common factors (typically much smaller than k),
- $F_1, F_2, \ldots, F_m$ are assumed to have zero means and unit variances,
- l_{ij} is called the *loading* of the ith variable on the jth common factor,
- $e_1, e_2, \ldots, e_k$ are independent of one another and the common factors.

The above equations and assumptions constitute the *factor model*. Thus each variable is composed of a part due to common factors and a part due to its own unique factor. Comparing Equation 2 with Equation 1 we see that the factor model would be the mirror image of the principal components description if not for the presence of the e_i's and the requirement that $m < k$. Ideally, the number of factors should be known in advance.

A major implication of the factor model given above is that the variance of X_i is broken into two parts which together must add up to 1 because of standardisation.

- The *communality* h_i^2 of X_i is that part of the variance due to the common factors.
- The *specificity* u_i^2 of X_i is that part of the variance due to e_i.

The numerical aspects of factor analysis are concerned with finding estimates of the factor loadings (l_{ij}) and communalities (h_i^2). There are many ways of finding these quantities, the process is called *initial factor extraction*. One method which is used in this paper is to use the major principal components found in PCA as the initial factors. As was shown in Section 3.2.2 that the initial factors (i.e. the first two principal components in this case) are not all that easy to interpret, this is true no matter what method is used for initial factor extraction. Consequently in FA the variables are transformed (rotated) in the factor space to change the loadings so that they are either near zero or near ± 1. Ideally we wish FA to give, for any given variable, a high loading on only one factor. If this is the case, it is easy to give each factor an interpretation arising from the variables with which it is highly correlated (high loadings). The rotations can be *orthogonal*, that is the rotated factors remain uncorrelated or *oblique*, the resultant factors are correlated. Oblique rotations are useful if for the sake of clearer interpretation we are prepared to relax the requirement that the factors are uncorrelated.

3.3.2 Results of FA on PRAC data

On the basis of the PCA and in order to keep the analysis simple, two factors were specified. The following table contains factor loadings for the unrotated, orthogonal and oblique factor solutions. The numbers in the F_1 and F_2 columns are the loadings l_{i1} and l_{i2} values (see Equation 2). The following approach is taken from Dillon and Goldstein (1984) p69.

1. For each variable, place an asterisk against the loading with the largest absolute value.

2. Then underline the loading if it is of *practical significance*, which in this case means that $l_{ij}^2 \geq 0.25$ (the loading is the correlation of the ith variable with the jth factor, hence the square of the loading will indicate what proportion of the variable's total variation is accounted for by factor j, so l_{ij} values less than 0.5 are considered of little practical importance).

3. Try to assign some meaning to the pattern of factor loadings. Assign a name or label that reflects to the greatest extent possible the combined meaning of variables that load on each factor.

Variable X_i	Description (See Appendix I)	Unrotated F_1	Unrotated F_2	Orthogonal F_1	Orthogonal F_2	Oblique F_1	Oblique F_2
Q1	'.. is directly applicable to identified .. R&D goal'	.71*	-.39	.24	.77*	.05	.78*
Q2	'.. significant scientific merit'	.78*	.31	.77*	.32	.79*	.09
Q3	'.. objectives clearly stated'	.75*	.22	.69*	.36	.69*	.15
Q4	'.. understanding of topic indicated'	.76*	.31	.76*	.31	.78*	.07
Q5	'.. methods .. appropriate'	.81*	.25	.75*	.38	.75*	.16
Q6	'.. personnel/facilities adequate for project'	.74*	.02	.55*	.50	.48*	.37
Q7	'.. costs reasonable in relation to .. benefits'	.70*	-.21	.36	.64*	.23	.58*
Q8	'.. increases in productivity to woolgrowers..'	.67*	-.32	.26	.69*	.10	.68*
Q9	'.. results .. applicable to entire .. industry'	.16	.27*	.31*	-.08	.37*	-.20
Q10	'.. duplication ? .. warranted ?'	.45	.50*	.67*	-.05	.78*	-.29
Q11	'Communication to .. users .. easily achieved'	.58*	-.51	.06	.77*	-.15	.84*
Q12	'.. short time before findings .. adopted..'	.55	-.63*	-.05	.83*	-.29	.95*
Q13	'.. probability of achieving project objectives..'	.73*	.0	.52*	.51	.45*	.39
Q14	Overall Project Rating	.88*	.12	.71*	.52	.66*	.34
Q15	Project Ranking	.58*	.08	.48*	.34	.44*	.22

In the unrotated factor solution, many of the *supply* (see PCA discussion above) variables load heavily on F_1 (Q2 to Q5, Q14) whereas some of the *demand* variables (Q11, Q12) have appreciable loadings on F_2. Although there seems to be some pattern it is not clear cut. If we imagine the loadings are like the co-ordinates of the variables in F_1 and F_2 space then rotation of the axes may increase (or decrease) the variable loadings thus making the factors easier to interpret. The orthogonal rotation creates new factor axes which are still at right angles to one another as were the original axes. The above table shows that the orthogonal solution is more clear cut: variables Q2-Q6, Q10, Q13-Q15 clearly load on F_1 and Q1,Q7,Q8,Q11 and Q12 clearly load on F_2. If a rotation of factor axes is used that does not require the new axes to be orthogonal (oblique rotation) the pattern is further clarified. The effect of the oblique rotation is to make the non-significant factor loading closer to zero.

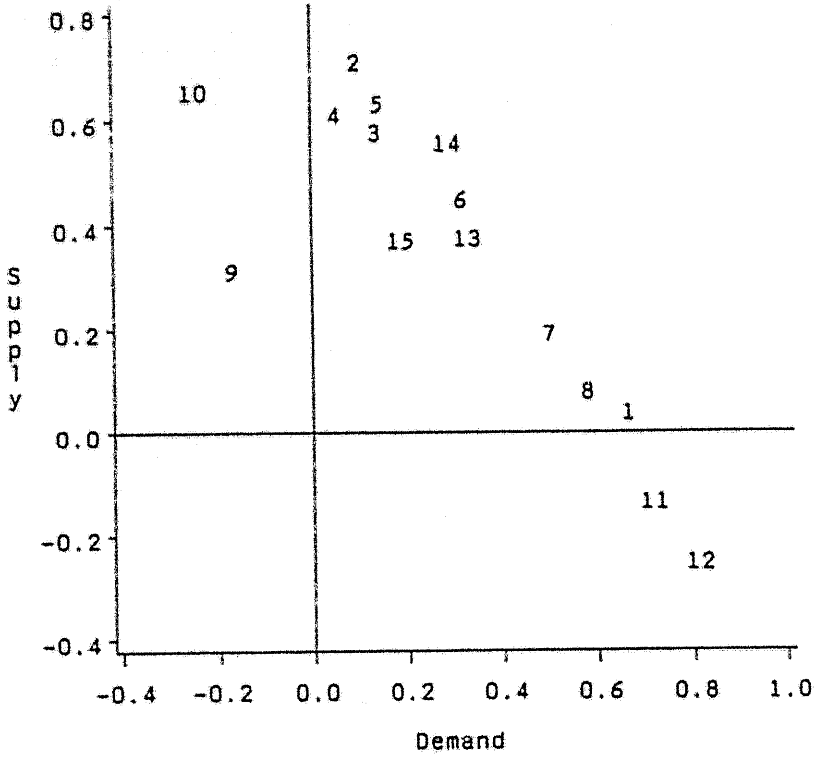


Figure 4: Correlations of Q1-Q15 with
Oblique Factors (Reference Structure)

Figure 4 is the graphical representation of the results of the oblique rotation where F_1 has been renamed 'Supply' and F_2 renamed 'Demand'. It is a plot of the *reference structure* of the factor solution which is the matrix of semipartial correlations between the variables and the common factors of 'Supply' and 'Demand', where for each factor the effect of the other factor has been removed. The use of the reference structure arises because with an oblique rotation the common factors are correlated so that the variable loadings do not give a clear indication of the simple correlations of the variables with the factors (for orthogonal rotations the factor pattern and the reference structure are equal). Notice the variables grouped together high on the F_1 axis (Q2, Q3, Q4, Q5 and Q14) and those associated with the F_2 axis (Q1, Q8 and to a lesser extent Q11, Q12, Q7).

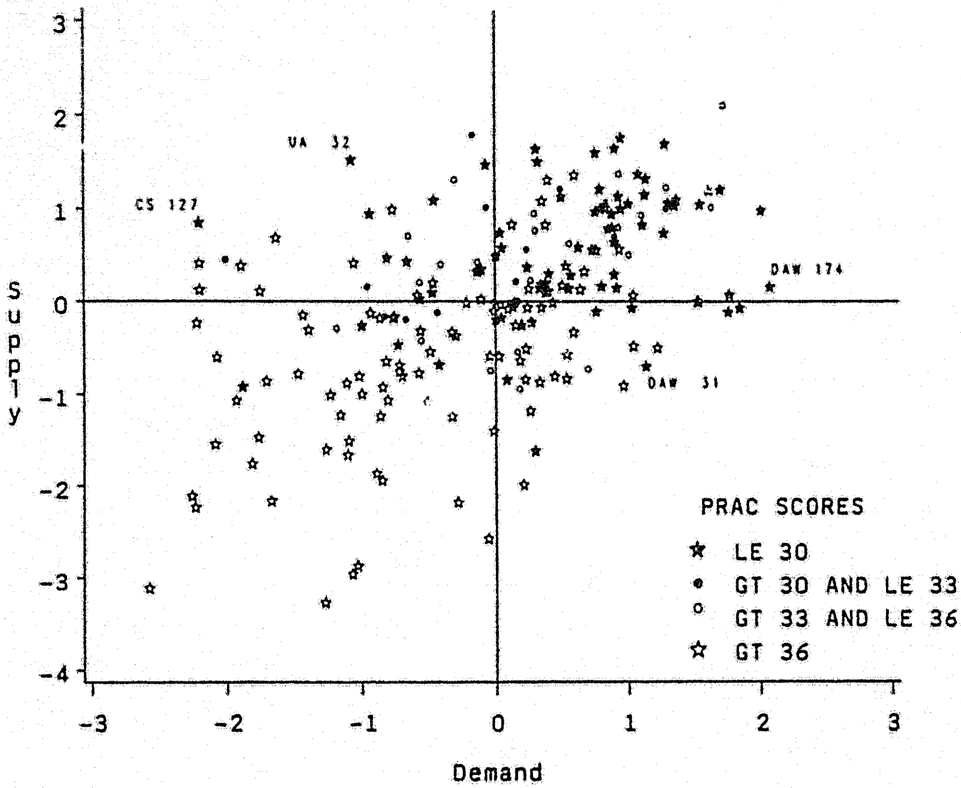


Figure 5: Plot of Supply and Demand Factors

Figure 5 is a plot of the factor scores. It is instructive to compare this plot with the principal components scores (Figure 3) to see how the two techniques differ but give complementary results. In the PCA almost 60% of the total variance was described by the first two principal components. Y_1 seemed to be associated with *supply* questions while Y_2 seemed to contrast *supply* questions with *demand* questions. FA on the other hand pesupposes the existence of underlying factors (we chose 2 in the case of the PRAC data because of PCA results) that generate that part of the variables' variation that it shares with the other variables. In PCA p orthogonal principal components are needed to completely describe the vrriance of the p variables and it is hoped that only the first few components account for the majority of the variance. The need for p components and the orthogonality constraint means that the principal components are not always easy to interpret. In FA q ($q < p$) factors are assumed to generate the covariance structure of the p variables. As q is often considerably smaller than p and factors need not be orthogonal the factors may be manipulted by various rotations to improve their interpretability. In the case of the PRAC data FA has supported the idea suggested by PCA that the referee questions are really measuring the latent *supply* and *demand* factors and that the success or otherwise of various proposals is more simply and informatively seen as the interaction between these two factors. In Figure 5 F_2 is identified as 'Demand' whereas in Figure 3 Y_2 is identified with the *difference* between supply and demand factors. Consequently large negative 'Demand' values correspond to high Y_2 values.

4 Standard Regression Methods and Extensions

The regression methods described in this section initially attempt to model the PRAC score as a function of proposal cost, proposal source (a dummy variable: 1=CSIRO and 0=All others) and the 15 referee question scores. The model is an example of a *multiple linear regression* because it regresses many *explanatory* or *predictor* variables (cost, source and referee question scores) on a single *dependent* variable (PRAC score) and is *linear* in its parameters. The meaning of this linearity property will become clearer when the model is mathematically formulated below. Apart from its simplicity this type of regression model is adopted because without other knowledge it may reasonably approximate how the PRAC uses the data. Essentially, the model assumes that the data relating to referee question scores, cost or source are weighted by their importance to the committee and combined in an additive fashion. Hence the model parameters are interpreted as a measure of the value given to a particular explanatory variable. More complex models could be suggested (i.e. include product terms to account for interaction between variables) but to justify such models would require more information about the scoring process and would complicate the interpretation of results.

4.1 Results of Regression Methods

Both the standard and truncated models fit the same linear equation to the data;

$$PRAC_i = \beta_0 + \sum_{j=1}^{n=15} \beta_j q_{ij} + \beta_{16} CS_i + \beta_{17} cost_i + \varepsilon_i \tag{3}$$

where;

$$\begin{aligned}
PRAC_i &= \text{PRAC score } i\text{th proposal},\\
\beta_0 &= \text{intercept parameter},\\
\beta_j &= \text{parameter associated with } j\text{th referee score},\\
q_{ij} &= i\text{th proposal's } j\text{th referee's score},\\
\beta_{16} &= \text{proposal source parameter},\\
CS_i &= i\text{th source dummy variable (CSIRO:CS=1, otherwise:CS=0)},\\
\beta_{17} &= \text{cost parameter},\\
cost_i &= \text{cost of } i\text{th proposal},\\
\varepsilon_i &= \text{error of model for the } i\text{th observation}.
\end{aligned}$$

The values that the PRAC score can take are bounded (i.e.$12 \leq$ PRAC ≤ 48) hence the application of OLS results in biased parameter estimates. This bias results in the OLS model overestimating low PRAC scores and underestimating high score (see figure 6). A truncated regression was carried out on the PRAC data and the results presented with OLS results.

Variable	OLS Estimates			Truncated Estimates		
	Coefficient	T-ratio	(Sig.Lvl)	Coefficient	T-ratio	(Sig.Lvl)
INTERCEPT	68.2	14.3	(.000)	96.3	8.99	(.000)
Q1	-.420	-.406	(.685)	.54E-01	.030	(.976)
Q2	.691	.697	(.486)	.749	.442	(.659)
Q3	1.53	1.52	(.126)	2.13	1.11	(.266)
Q4	-3.18	-3.17	(.002)*	-4.94	-2.47	(.013)*
Q5	1.48	1.33	(.185)	2.01	.979	(.327)
Q6	-2.05	-1.90	(.057)*	-4.17	-2.05	(.040)*
Q7	-2.22	-2.54	(.011)*	-3.76	-2.69	(.007)*
Q8	-1.07	-1.10	(.270)	-2.71	-1.84	(.065)*
Q9	1.26	2.27	(.023)*	1.99	2.17	(.030)*
Q10	-.822	-1.40	(.161)	-1.86	-1.91	(.056)*
Q11	-1.42	-1.47	(.140)	-2.43	-1.53	(.127)
Q12	1.12	1.20	(.229)	2.33	1.52	(.129)
Q13	-5.49	-1.36	(.174)	-7.12	-1.01	(.313)
Q14	-2.78	-2.02	(.043)*	-3.20	-1.29	(.196)
Q15	-.770	-1.43	(.150)	-1.06	-1.32	(.187)
CS	-2.84	-1.93	(.054)*	-4.09	-1.79	(.073)*
COST	.303E-05	.206	(.837)	.740E-05	.309	(.758)
MSE	7.20			8.50		
R^2_{adj}	0.48			NA		

The fit of the two models to the data on the basis of the mean square error (MSE) and the adjusted R^2 were judged to be only fair. The model accounts for about half of the observed variation in the sample. Whether this is satisfactory depends on the aims of the analysis. If the aim is to examine the relative importance of explanatory variables then such a fit is adequate. However if the aim is to use the model for prediction then the fit is inadequate, there is too much unexplained variation to be confident about model predictions. The truncated model has a larger MSE implying that it accounts for less variation than the OLS model (the behaviour of R^2 is more complex with iterative methods so it is not calculated).

All significant coefficients, bar one, had the expected sign. On the basis of this analysis, key variables appear to be Q4, Q6, Q7 and Q14. Also the dummy variable indicating CSIRO proposals is marginally significant. This implies, given the same score for questions 1 to 15, that proposals from CSIRO are favoured over those coming from other research institutions.

The 'wrong' sign on Q9 implies that projects which are 'regional specific' tend to be favoured over projects whose results will be generally applicable to the entire Australian wool industry. This counter intuitive result occurs because there appears to be a negative association between Q9 and all the other variables which affect the decision to fund a proposal. In other words, the projects which are regional specific tend to be better defined, relevant and more manageable than those which might be generally applicable across the whole wool industry.

Actual scores and predicted scores from both OLS and truncated models are plotted against project ranked by PRAC score and OLS score. This results in a step like curve for actual scores and irregular saw tooth curves for the predictions. Note that the truncated model produces a curve that is a more exaggerated version of the OLS model and that it tracks the actual values more closely although it does this with a greater spread of values.

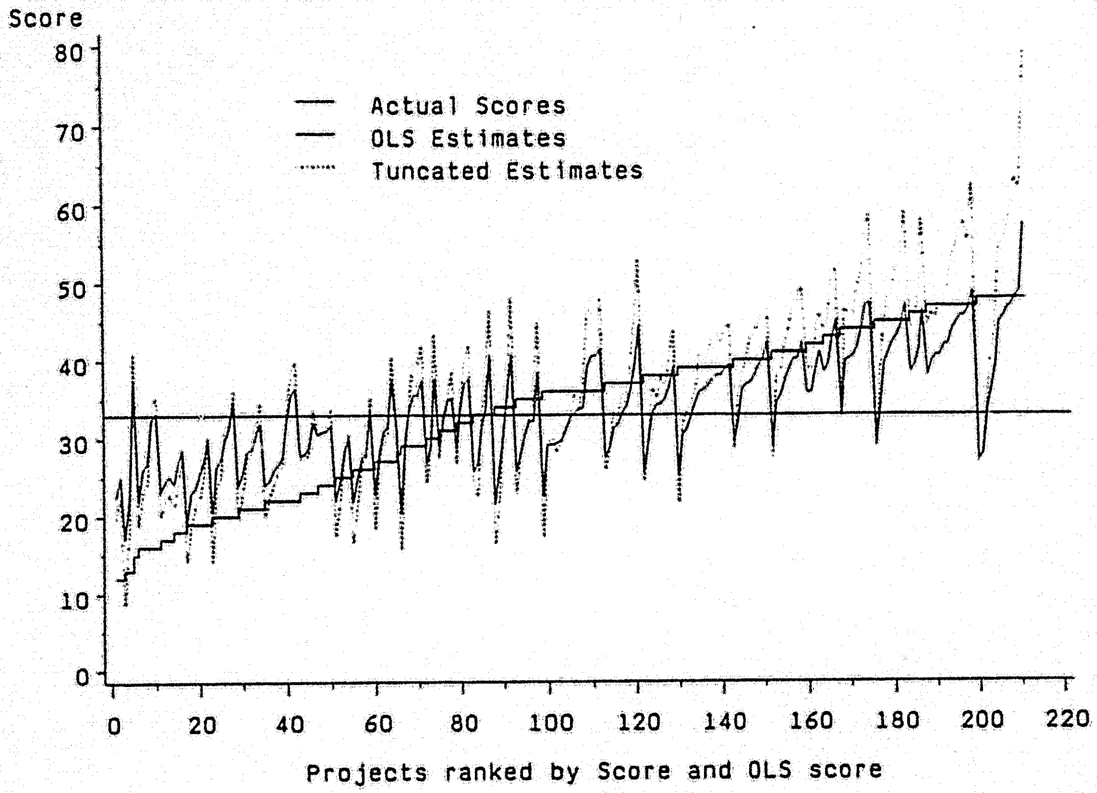


Figure 6: Comparison of Regression Predictions and Actual Scores

An alternative method of examining the effectiveness of the OLS model is to tabulate the number of wrong predictions.

Project Grouping (By PRAC score)	Number of Proposals in Sample	Number Incorrect	Percent Incorrect
If Score < 29	65	7	11
If 29 ≤ Score ≤ 31	64	36	56
If Score > 38	82	11	13
TOTAL SAMPLE	211	54	25

The model's predictions reported in the above table suggests that there still exists a fair amount of unexplained variation between the explanatory variables and the PRAC decision. The model does poorly in predicting the PRAC decisions for marginal projects but performs quite well for projects with high or low scores. Differences in the referee's ranking and the committee score may occur because of one or more of the following reasons.

- The PRAC disagrees with the referee's responses to questions 1 to 15 in the referee reports. Increasing the number of referees per proposal may increase the correlation between referee scores and the PRAC decision.

- The empirical analysis assumes that the PRAC treats all proposals independently.
- The referee's comments at the bottom of the referee's reports are not used in modelling the scores.
- Other factors could have influenced the PRAC's decisions as evidenced by the inclusion of the CSIRO dummy variable in the model that turned out to be significant.

5 Qualitative Response Models (QR)

An alternative to modelling PRAC scores is to model the decision to fund or not to fund. This requires the creation of a *binary* dependent variable which takes the value 1 if the proposal is funded and 0 if it is not. Such a dependent variable is described as categorical or qualitative because it is really a numerical code for some attribute of the data rather than a measurement. Models that use categorical or qualitative dependent variables are called qualitative response (QR) models (see Amemiya (1981),(1985) and Maddala (1983)). The simple binary categorisation of the PRAC decisions is not the only one possible and more sophisticated schemes identifying high, middle and low scoring projects could be described. In addition there is a brief note about the relationship of *discriminant analysis* to QR models because procedures for doing discriminant analysis are more commonly available and the results are easier to explain than those of QR models.

5.1 Linear Probability Model

The linear probability model represents a direct application of classical regression techniques to modelling a QR variable;

$$q_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_j x_{ij} + \ldots + \beta_k x_{ik} + \varepsilon_i \tag{4}$$

where q_i is the ith observation on qualitative variable. The assumptions remain the same as for OLS.

Although the linear probability model is algebraically similar to the standard regression model its interpretation is fundamentally different. The standard model is a model for the response of Y; given a set of explanatory variable observations it provides an estimate for the expected value or level of Y. In the QR model the response level of Y is not at issue (in the simplest case; 0 or 1) what is at issue is which level is likely to occur. Hence QR models are not models of the level of response of Y but models for the *probability* that Y is either 0 or 1 for instance. This difference is not all that surprising when one considers that QR models are describing the occurence of a qualitative attribute of the data (success or failure, alive or dead etc) not a continuously varying quantitative response. The QR dependent variable is no more than a numeric code for these attributes.

The linear probability model has problems in that the dependent variable is not constrained between 0 and 1. There is nothing in the model to stop extreme values of the explanatory variables producing predicted values of the dependent variable greater than 1 or less than zero.

5.2 The Logistic Model

As noted above the linear probability model is inconsistent in its assumptions and does not constrain the probabilities it is estimating to fall between 0 and 1. An approach that overcomes these difficulties is the logit model which is defined as the natural logarithmic value of the odds in favour of a positive response, that is

$$\mathrm{L}_i = \ln\left(\frac{\pi_i}{1-\pi_i}\right) = \mathbf{x}_i'\beta \tag{5}$$

where π_i is the conditional probability of a positive response for the proposal with characteristic $\mathbf{x}_i'$ and the β is vector of parameters. It follows that

$$\pi_i = \frac{1}{1 + e^{-\mathbf{x}_i'\beta}} \tag{6}$$

In the case of the logit model the coefficients reflect the effect of a change in the regressor variable on $\ln(\pi_i/(1-\pi_i))$. The amount of the increase depends on the original probability and thus on the initial values of all the independent variables and their coefficients.

5.3 Results and Conclusions of QR Models

	Linear Probability Model			Logistic Model		
Variable	Coefficient	T-Ratio	(Sig.Lvl)	Coefficient	CHI-SQUARE	(Sig.Lvl)
INTERCEP	-0.841	-3.07	(0.003)	-9.80	20.2	(0.000)
Q1	0.029	0.491	(0.624)	.034	0.01	(0.932)
Q2	-0.020	-0.343	(0.732)	-.078	0.05	(0.832)
Q3	-0.132	-2.284	(0.024)*	-.889	5.02	(0.025)*
Q4	0.162	2.796	(0.006)*	1.029	6.90	(0.009)*
Q5	-0.072	-1.12	(0.266)	-.370	0.80	(0.372)
Q6	0.057	0.919	(0.359)	0.499	1.25	(0.264)
Q7	0.127	2.53	(0.012)*	0.943	7.35	(0.007)*
Q8	0.043	0.769	(0.443)	0.520	2.04	(0.153)
Q9	-0.051	-1.60	(0.111)*	-.[illegible]61	3.10	(0.078)*
Q10	0.0366	1.09	(0.279)	0.348	2.21	(0.138)
Q11	0.0445	0.801	(0.424)	0.311	0.75	(0.386)
Q12	-0.0787	-1.47	(0.142)	-.668	3.43	(0.064)*
Q13	0.290	1.25	(0.214)	1.468	0.83	(0.363)
Q14	0.120	1.52	(0.131)	0.658	1.73	(0.188)
Q15	0.023	0.738	(0.461)	0.119	0.43	(0.513)
CS	0.159	1.88	(0.062)*	0.987	3.52	(0.061)*
COST	2.47E-07	0.292	(0.771)	3.2E-06	0.28	(0.596)

The results of the OLS estimates of the linear probability model and the maximum likelihood estimates of the logit model are presented in the above table. The R^2_{adj} for the linear probability model is 0.30, showing that the model performs worse than the OLS model. However this is to be expected as we are trying to fit a multiple linear regression to a binary variable. In logistic regression there are many methods for assessing goodness of fit (see Amemiya, 1981), one that is superficially similar to the OLS R^2 is McFadden's $R^2 = 1 - l/l_0$, where l is the log likelihood after fitting the full model and l_0 is the log likelihood of fitting only the intercept term. This can be approximately calculated from the output and is 0.27. Hensher and Johnson (1981) state that values between 0.2 and 0.4 are considered extremely good fits. This model predicts about 75% of the committee's decisions.

As with the OLS results using the PRAC scores, most coefficients have the expected sign (signs are opposite to those of the regression models because successful projects have a higher dependent variable value than unsuccessful projects in QR models) except on Q3, Q9 and Q12. The coefficient on Q3 suggests (erroneously) that funded projects do not necessarily have their objectives more clearly stated than those proposals which are not funded. In respect to the coefficient on Q9 see previous discussion. The coefficient on Q12 implies that proposals which have a long adoption phase are preferred by the committee; this is unlikely. It is possible however that the significant coefficient on Q12, at least for the logit model, reflects the committee's desire to fund good basic research which naturally has associated with it long adoption lags. In summary the results of the QR models confirm the earlier analyses.

6 Discriminant Analysis (DA)

> Discriminant techniques are used to classify individuals into one of two or more alternative groups (or populations) on the basis of a set of measurements. The populations are known to be distinct, and each individual belongs to one of them. These techniques can also be used to identify which variables contribute to making the classification. (Afifi and Clark (1984) p24)

The reason for discussing DA in this paper is that it is very closely related to logistic regression both conceptually and algebraically. Furthermore by tackling the problem from the viewpoint of discrimination or classification it is possibly easier to interpret and to explain to non-statisticians than a regression of the log of the odds. Finally it introduces concepts and methods of *Bayesian* statistical inference which unlike the inference discussed so far in this paper (described as *frequentist* or *classical* inference) allows the use of *prior* knowledge in the analysis of the problem which is quite independent of that knowledge gained from the sample. The classical approach only allows analysis on the basis of a gathered sample of data and does not give the analyst a formal way of using prior knowledge which may have some bearing on the results.

To simplify the discussion consider two groups or populations g_0, g_1. Let the groups be distinguished by a binary variable Q, i.e. $Q = 0$ for g_0 and $Q = 1$ for g_1. The task that DA sets itself is to decide the probable membership of the ith individual using measurements on a set of k variables $\mathbf{x}_i' = (x_{i1}, \ldots, x_{ik})$. The results of DA analysis can be used two ways, firstly it can be used to assess the importance of the variables in distinguishing between the groups and secondly it can give probable group membership to as yet an unallocated individual. Symbolically we want to compute,

$$\pi_{i1} = \Pr(Q = 1|\mathbf{x}_i') \tag{7}$$

i.e. the probability that the individual belongs to the population distributed as g_1 *given* or on the basis of the set of measurements $\mathbf{x}_i'$. Note that in two group case probability of membership of g_0 is $\pi_{i0} = 1 - \pi_{i1}$. The following assumptions are made.

- The probability distributions corresponding to g_1 and g_0 are known or can be estimated. These allow the conditional probabilities $\Pr(\mathbf{x}_i'|Q = 1)$ and $\Pr(\mathbf{x}_i'|Q = 0)$ to be computed.
- There is a *prior probability* for group membership: $\Pr(Q = 1) = p_1$ and $\Pr(Q = 0) = p_0$. If these are unknown then they can be both set to 0.5 so that *before* measurement the individual is equally likely to belong to either group. Alternatively they can be proportions of group size to combined group size thus reflect the greater likelihood of an individual being in the larger group.

In order to compute π_{i1} (called the *posterior probability*) from the known or estimated conditional probabilities $\Pr(\mathbf{x}_i'|Q = 1)$ and $\Pr(\mathbf{x}_i'|Q = 0)$, *Bayes' Theorem* is used.

$$\pi_{i1} = \Pr(Q = 1|\mathbf{x}_i') = \frac{\Pr(\mathbf{x}_i'|Q = 1)\Pr(Q = 1)}{\Pr(\mathbf{x}_i'|Q = 1)\Pr(Q = 1) + \Pr(\mathbf{x}_i'|Q = 0)\Pr(Q = 0)} \tag{8}$$

or

$$\pi_{i1} = \Pr(Q = 1|\mathbf{x}_i') = \frac{\Pr(\mathbf{x}_i'|Q = 1)p_1}{\Pr(\mathbf{x}_i'|Q = 1)p_1 + \Pr(\mathbf{x}_i'|Q = 0)p_0} \tag{9}$$

In Amemiya (1981 p1508), QR models and DA are contrasted;

In econometric and biometric QR models, the determination of $\mathbf{x}_i'$ (e.g. income or dosage) clearly precedes that of Y_i (e.g. purchase or death); therefore it is important to specify $\Pr(Y = 1|\mathbf{x}_i')$ whereas the distribution of $\mathbf{x}_i'$ may be ignored. On the contrary, in the DA model, the statement $y_i = 1$ (e.g. a skull belongs to a man) logically preceeds the determination of $\mathbf{x}_i'$ (skull measurements); therefore it is more important to specify the conditional distribution of $\mathbf{x}_i'$ given y_i.

To use Equation 9 it is usually assumed that the conditional distributions are *multivariate normal.* To more easily see how this assumption allows the simple calculation of posterior probabilities from a logistic like expression consider $\mathbf{x}_i' = x_i$ (i.e. only one explanatory variable measured on each individual so that the conditional distributions are univariate normal):

$$f_{g_1} = \Pr(\mathbf{x}_i'|Q = 1) = \mathrm{N}(\mu_1, \sigma_1^2) = (1/\sqrt{2\pi\sigma_1^2})\exp[-(x_i - \mu_1)^2/2\sigma_1^2] \tag{10}$$

and

$$f_{g_0} = \Pr(\mathbf{x}_i'|Q = 0) = \mathrm{N}(\mu_0, \sigma_0^2) = (1/\sqrt{2\pi\sigma_0^2})\exp[-(x_i - \mu_0)^2/2\sigma_0^2] \tag{11}$$

Substituting these distributions into Equation 9 we get an expression that is similar to Equation 6:

$$\pi_i = \frac{1}{1 + e^{-\mathrm{L}(x)}} \tag{12}$$

where $\mathrm{L}(x)$ is called the *discriminant function* and is a function of $\mu_1, \mu_0, \sigma_1, \sigma_0$ and the logarithm of the prior probabilities.

6.1 Results of DA

One way of assessing the performance of DA is the *error rate* (ϵ). Note prior probabilities estimates of the observed proportions of unsuccessful and successful proposals were 0.5877 and 0.4123, respectively. The error rate is the expected proportion of cases missclassified by the DA classification rule. There are a number of methods for estimating ϵ the most common being the *apparent error rate* ($\hat{\epsilon}_A$) which is simply the ratio of the number of missclassifications (e) over total sample size (N); $\hat{\epsilon}_A = e/N$. In the following two-way tabulation of actual observations against DA classification we see that ninety-six or 77 % of the unsuccessful proposals were classified correctly while approximately 83 % of the successful proposals were classified correctly.

		Predicted		Total
		Not Funded (%)	Funded (%)	
Actual	Not Funded (%)	96 (77.4)	28 (22.6)	124 (58.8)
	Funded (%)	15 (17.1)	72 (82.8)	87 (41.2)
Total (%)		111 (52.6)	100 (47.4)	211 (100)

This gives an $\hat{\epsilon}_A$ of $(28 + 15)/211 = .204$. Unfortunately this estimate is biased low (i.e. an optimistic estimate of ϵ). This is to be expected since the estimate of ϵ has been derived from the sample uced to generate the classification rule. Ideally ϵ should be estimated using data from another sample to get a true idea of the success of classification. However the bias decreases with sample size and hence for large samples $\hat{\epsilon}_A$ is a useful estimate of ϵ. By 'large' it is meant the sample has more cases than than ten times the number of variables (James, 1985) which is true of the PRAC data. An important problem with PRAC data set is that groups within it are unlikely to have multivariate normal distributions. This is because many of the variables are ordinal or binary rather than continuous. In addition no attempt has been made to assess the importance of outliers in the data, the presence of which could effect discrimination by their influence on the sample estimates of group means and variances. However the degree of successful discriminantion implies that failure of the PRAC data to meet relevant distributional assumptions does not completely compromise the use of DA. In terms of the study aims the DA usefully shows that the quantitative information on proposals does indicate proposal funding status in 80 % of the cases. Clearly if more use of this technique were to be made then more care would have to taken but at this stage it is sufficient to show that it performs well in comparison to other techniques and is possible a useful way of modelling the PRAC decision making process.

One way to approach some of the distributional problems outlined above is to use a three way grouping of proposals on the basis of PRAC scores; highly successful (1, PRAC score 12 to 29), marginal (2, PRAC score 30 to 37) and highly unsuccessful (3, PRAC score 38 to 48). One would expect the distribution of data within these groups to be less skewed than in the binary case where highly successful or unsuccessful proposals may appear as outliers.

		Predicted			Total
		'Good' (%)	'Marginal' (%)	'Bad' (%)	
Actual	'Good' (%)	61 (85.9)	7 (9.9)	3 (4.2)	71 (33.7)
	'Marginal' (%)	2 (4)	42 (84)	6 (12)	50 (23.7)
	'Bad' (%)	8 (8.9)	8 (8.9)	74 (82.2)	90 (42.7)
Total (%)		71 (33.7)	57 (27)	83 (39)	211 (100)

This gives an $\hat{\epsilon}_A$ of $(7+3+2+6+8+8)/211 = .161$.

In a similar fashion to stepwise regression, stepwise discriminant analysis can be implemented which selects that subset of variables that contribute most to the successful discrimination (see Afifi and Azen, 1979 for more details). In the case of the PRAC data we get the following results;

Variable	Partial R**2	F Statistic	Prob > F
Q14	0.2329	63.450	0.0001
Q7	0.0402	8.711	0.0035
Q4	0.0239	5.070	0.0254
CS	0.0242	5.110	0.0248
Q3	0.0235	4.938	0.0274
Q9	0.0163	3.381	0.0674

6.2 QR models and DA compared

In the following graph the probability of success that the QR models predict for each proposal is plotted. For comparison, the actual PRAC decisions are also plotted where a project that was successful is given a probability of sucess of one and a project that was unsuccessful is given a probability of zero. The unconstrained nature of the Linear Probability model is immediately obvious with some probablities greater than one or less than zero. However accepting these problems with the model and considering the noisy nature of the data the Linear Probability model does not fair too badly (see Byron (1988) for further empirical examples). The graph clearly demonstrates the superiority of DA over the logistic model which is due to the former's use of prior information where prior probabilties were estimated by the proportions of successes and failures in sample.

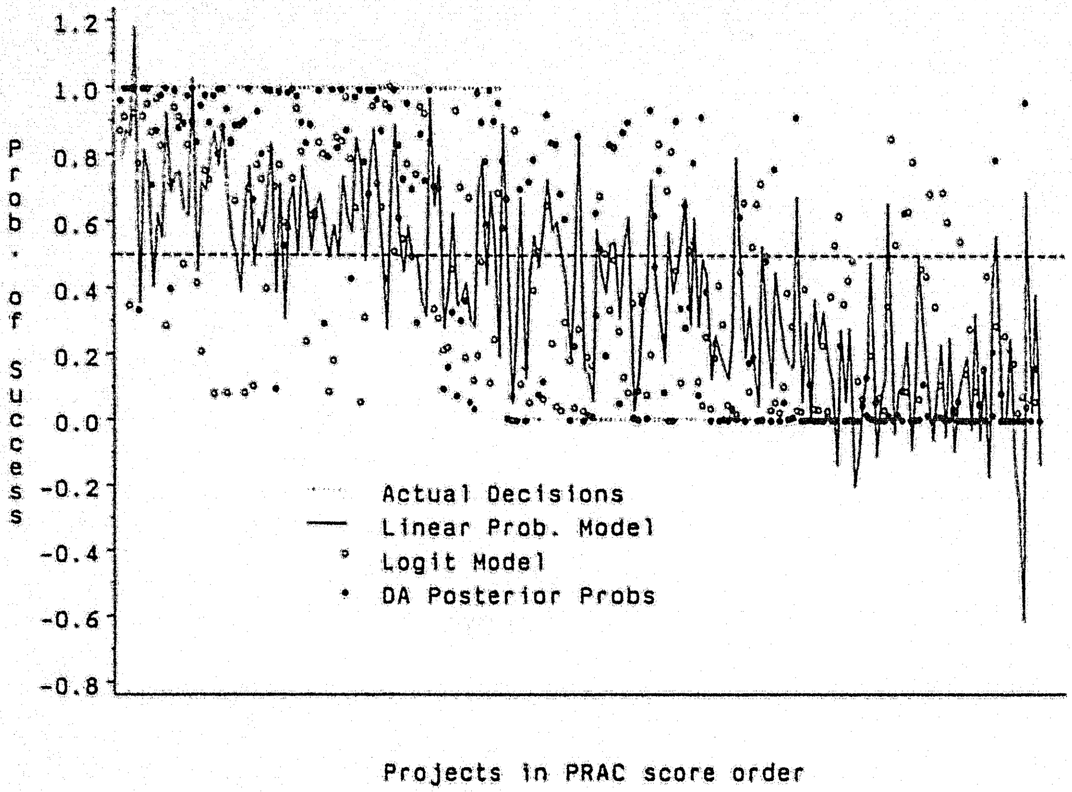


Figure 7: Graphical Comparison of QR Models and DA

7 Summary and Conclusions

As mentioned in the introduction the major aim of the study is to examine the use made of the 'quantitative' information available to PRAC. The answers to questions such as: How (if at all) is this information used by the committee? If it is used which is the most influential? Can unusual decisions be detected? Can the data be usefully summarised without sacrificing important information? etc. Such aims are necessarily vague and limited to the structure found in the data because we have little information as to how the committee members reach decisions. We only have the final outcome of their deliberations and quantitative information they may have used for each proposal. Bearing this in mind the techniques used in this study will now be assessed.

Due to the complex nature of the PRAC data simple **histograms** and **scattergrams** are of little value on their own because they do not display the *multivariate* structure of the data.

The **sample correlation matrix** in an easy to read format proved an effective way of examining the structure of the data in addition to being an initial stage of several other methods. The matrix revealed high correlations between the explanatory variables derived from the referee questionnaire and relatively low correlations with the PRAC score (dependent variable) and other explanatory variables such as cost and the CSIRO dummy variable. These findings suggest that models trying to relate decisions to the explanatory variables will find the data 'noisy'. They also suggest that an examination of the referee data may reveal interesting structure.

In order to examine the referee data further and possibly summarise its content **Principal Components Analysis (PCA)** and **Factor Analysis (FA)** were used. Both techniques re-express existing variables as new variables called principal components or fac ... that account for the variance and covariance structure of the data, respectively. Both techniques are valuable when very few principal components or factors are required (*dimensional reduction*) and when these new variables have informative interpretations in the context of the study. Both techniques when applied to the PRAC data supported the idea that the variation exhibited by the 15 variables derived from the referee questionnaire is largely an expression of two factors. One factor called 'Supply' measures the quality of available scientific expertise and resources that a proposal may have and the other factor called 'Demand' measures the extent to which the expected results of the proposal may benefit or be applied to the industry. Not only is this an interesting finding in itself but the two factors provide a convenient tool for graphically looking at trends in the PRAC decisions as a whole. In particular, examination of principal component and factor score plots, and some of their 'outliers' revealed that the PRAC tended to approve proposals more on the basis of the supply 'factor' than of the 'demand' factor. This 'bias' may arise because the committee members may feel more confident about assessing the scientific value or merit of a proposal rather than the extent to which it meets industry needs. In order to draw the attention of both the referees and the PRAC to these 'supply' and 'demand' factors the referee quetionnaire was redrafted (see Appendix I and II).

So far the methods used on the PRAC data have been explorato... (in this study FA is considered an exploratory technique), that is they have been aimed at revealin... the structure of the present data set without making any assumptions about how the data is generated or making a distinction between explanatory or dependent variables. The rest of the methods that were applied to the PRAC data involved models that in various ways explicitly related explanatory variables to a dependent variable. Furthermore such models make assumptions about how the data values are distributed, allowing the analyst to carry out statistical tests that objectively assess the suitability of the model and generalise from the particular data set to the population of which it is seen to be a sample. Whether such modelling really advances the aims of the study will be discussed below.

The **Standard Regression Model** when fitted to the PRAC data assumes that for a given set of explanatory variable values the PRAC score has a normal distribution with a mean given by the weighted linear combination of the explanatory variables and a constant variance. Both the weights (co-efficients or model parameters) and the variance are estimated from the sample as a whole. Given the availability of an almost continously valued response variable such as the PRAC score, regression methods may seem a 'natural' or popular choice. Although the model did distinguish between influential and non-influential variables the overall fit of the model to the data was disappointing. This implies that a simple linear model does not account for the majority of the observed variation. This is not surprising when one considers that the PRAC score is the sum of the 12 committee member scores. These scores may not be arrived at independently of one another and may not treat all input data independently (i.e. the model does not allow for interactions between variables). Proposals are no doubt assessed in relation to the need for research within certain well defined areas. Research areas or types were not incorporated into the model thus leaving out factors that may have may influenced the scores. For all these reasons the level of the PRAC score may not be a simple linear combination of the available explanatory variables thus limiting the usefulness of the standard regression model in this study.

Quite apart form the *model specification problems* discussed above is the failure of the PRAC data to meet distributional assumptions. Possible lack of independence between proposals has already been mentioned but another problem was the truncated nature of the PRAC data. As explained this leads to biased model parameter estimates as evidenced by the over and underestimating of low and high PRAC scores respectively (see Fig 6). A **Truncated Regression Model** was applied to the data to see if allowance for this property of the PRAC scores made much difference. Overall there was not a great difference between the models however in the truncated model the variance was greater and the parameter estimates generally less significant.

The **linear probability model** and the **logit model** are examples of *qualitative response models* where it is the probability of success or failure of a proposal to gain funding that is modelled. The linear probability model was introduced because it is easy to implement and to motivate the use of the logit model. It was found that the logit model fitted the data better than the regression models, this may be explained by observing that while there may be considerable variation in the actual PRAC score levels there is probably less variation in the actual committee decision to fund or not to fund, which after all is the outcome of most importance. The logit results were generally consistent with the other regression results but unfortunately are harder to communicate because the logit model gives the log of the odds ratio as a linear combination of the explanatory variables and requires iterative methods for solution.

Another technique which has a similar interpretation but may be easier to communicate is **Discriminant Analysis (DA)**. Rather than model the probability of success as a function of explanatory variables, DA classifies the observations into groups corresponding to success or failure on the basis of the explanatory variable values. DA produces a simple classification rule for deciding which group a proposal belongs to and can simply calculate a probability of group membership under certain distributional assumptions. It is simply generalised to more than two groups and allows for prior information about likely group membership to be incorporated into the model. From the point of view of communicating results it may be easier to view the problem as one of classifcation rather than one of regression. In fact the classificaion paradigm may be more appropriate to the work of the committee, that is the task is to assign proposals to a successful group or unsuccessful group (in terms of funding) rather than produce a formula for the level of the PRAC score. Seen this way the PRAC score is a means to an end and in itself a misleading measure of proposal merit. Furthermore DA can provide for variable selection methods similar to stepwise regression and feature selection methods which are similar to PCA (the later methods were not investigated in this study). The computational simplicity of DA rests on the assumption of multivariate normality which is unlikely to apply to the PRAC data. However the failure of this assumption largely effects the calculation of the posterior probability and does not effect the interpretation of the analysis. In the case of the PRAC data, of all the techniques, DA most successfully discriminated between funded and unfunded proposals so that greater attention to distributional assumptions (transformations, logistic discrimination etc) may have improved its performance.

What has the study revealed about the decisions of the Production Research Advisory Committee (PRAC)?

The quantitative data which is available to the PRAC does allow the prediction of proposal success with regard to funding. As this is unlikely to be purely coincidental it implies that the PRAC do use this data in its deliberations. Secondly the data derived from the Referee questionnaire can be seen as generated by 'Supply' and 'Demand' factors that relate to the quality of available research resources and the demand for its results. Finally the PRAC is 'biased' against, or less able to assess strong 'Demand' proposals compared to strong 'Supply' proposals.

Overall the study demonstates the importance of using a wide variety of techniques, the importance of dimensional reduction and the usefulness of high quality statistical graphics. It is envisaged that future work will progress along two paths. As new data becomes available the models developed in this paper will be used as a forecasting tool to rank projects from 'best' to 'worst' and help identify decisions at PRAC meetings which are inconsistent with referee assessments. Secondly, work may also be extended to examine the progress of funded proposals This would involve examining previously funded and 'completed' projects, developing a measure or measures of success of the projects and seeing what association exists, if any, between referee question scores 1 to 15 and project outcomes.

8 Appendix I: Old Referee Questionnaire

AUSTRALIAN WOOL CORPORATION | REFEREE'S REPORT - PRODUCTION RESEARCH | PROJECT NO: No

PROJECT TITLE: Title

Rating Procedure: (Circle Number)	Strongly Disagree	Disagree in General	Neither Agree nor Disagree	Agree in General	Strongly Agree
	1	2	3	4	5

GOAL: Goal

REFEREE NO:

	PROJECT PARAMETER	ASSESSMENT
1.	The project is directly applicable to the identified wool R&D goal	1 2 3 4 5
2.	The project has significant scientific merit	1 2 3 4 5
3.	Research objectives are clearly stated	1 2 3 4 5
4.	Thorough review of literature and understanding of the topic is indicated	1 2 3 4 5
5.	The methods proposed to complete the project are appropriate	1 2 3 4 5
6.	The personnel and facilities are adequate for the project	1 2 3 4 5
7.	Project costs are reasonable in relation to anticipated benefits	1 2 3 4 5
8.	Successful completion of the project will lead to significant increases in productivity for woolgrowers (i.e. same production/reduced costs or same costs/increased production)	1 2 3 4 5
9.	The project results will be generally applicable to the entire Australian wool industry (e.g. all geographical zones, wool types). If score 1 or 2 nominate the sector/area likely to benefit most	1 2 3 4 5
10.	Does this project duplicate past or present work (Y/N)? If yes, the duplication is warranted	1 2 3 4 5
11.	Communication of findings to potential users (researchers, advisory officers or woolgrowers) will be easily achieved	1 2 3 4 5
12.	It will be a short time before findings from the project will be adopted and benefits produced for woolgrowers	1 2 3 4 5
13.	The research has a probability of achieving the project objectives of about (0 = No possibility 1 = Certain achievement of objectives)	____
14.	OVERALL PROJECT RATING (Choose one) A or B or C or D (A = Excellent, B = Very Good, C = Good, D = Fair)	____
15.	PROJECT RANKING Rank this project in comparison with other projects you are refereeing	__/__

COMMENTS: Indicate the key reasons for your overall project rating and project ranking.

9 Appendix II: New Referee Questionnaire

REFEREE'S REPORT

RESEARCH & DEVELOPMENT GROUP

Australian Wool Corporation

TITLE	
	Project No.
	Goal No.
	Referee No.

Rating:	Strongly Disagree	Disagree in General	Neither Agree nor Disagree	Agree in General	Strongly Agree
Score	1	2	3	4	5

PROJECT ASSESSMENT CRITERIA	RATING (Circle a number)				
I. QUALITY OF RESEARCH PROPOSAL					
1. Research objectives are clearly stated.	1	2	3	4	5
2. Thorough review of literature and understanding of topic is indicated.	1	2	3	4	5
3. The methods pr.. .sed to complete the project are well described and appropriate.	1	2	3	4	5
4. The personnel and facilities are adequate for the project.	1	2	3	4	5
5. Project costs are reasonable.	1	2	3	4	5
6. Does this project duplicate past or present work (Y/N)?	☐				
If yes, the duplication is warranted.	1	2	3	4	5
II. BENEFITS TO WOOL INDUSTRY					
7. The applicant's assessment of the commercial sig-ificance of the research is realistic.	1	2	3	4	5
8. The project results will be generally applicable to the entire Australian wool industry (all geographical zones, wool types). If score 1 or 2 nominate the sector/area likely to benefit most.	1	2	3	4	5
9. The methods of technology transfer outlined will ensure efficient communication to potential users.	1	2	3	4	5
10. Would you anticipate that the findings from the project will be adopted and benefits produced for woolgrowers within (circle one).	5	10	15	20	years

III. ASSESSMENT

OVERALL PROJECT RATING: (circle one) fair good very good excellent

PROJECT RANKING: Rank this project in comparison with other projects you are refereeing. ____ / ____

COMMENTS: Indicate the key reasons for your overall project rating and project ranking. Typed comments preferred. Comments to applicants will be forwarded unedited and should be appended on a separate page.

RAD01-09/89

References

[1] Afifi, A.A. and Azen, S.P. (1979) *Statistical Analysis- A Computer Oriented Approach*, Academic Press.

[2] Afifi, A.A. and Clark, V. (1984) *Computer-Aided Multivariate Analysis*, Lifetime Learning Publications.

[3] Amemiya, T. (1981) "Qualitative Response models: A Survey", *Journal of Economic Literature*, **19**, pp 1483-1536.

[4] Amemiya, T. (1985) *Advanced Econometrics*, Basil Blackwell.

[5] Byron, R.P. (1988) *The Decision to Purchase a Suit in China: Some Results from the Nanjing Survay* ANU.

[6] Chatfield, C. and Collins, A.J. (1980) *Introduction to Multivariate Analysis*, Chapman and Hall.

[7] Dillon, W.R. and Goldstein, M. (1984) *Multivariate Analysis - Methods and Applications*, John Wiley.

[8] Hensher, D.A. and Johnson, L.W. (1981) *Applied Discrete Choice Modelling*, John Wiley

[9] James M. (1985) *Classification Algorithms*, Collins.

[10] Maddala, G.S. (1983) *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press.

[11] Vlastuin, C. (1989) "The Decision to Fund a Research Proposal: Some Results Using PRAC Data", Draft Paper, Australian Wool Corporation.