The Structure and Performance of Tasmanian Manufacturing Industries

by

Dangnoi Dana Hutchinson

submitted in partial fullfilment of the requirements

for the degree of

Masters of Financial Studies

University of Tasmania

November 1985

I certify that this thesis contains no material which has been accepted for the award of any other higher degree or graduate diploma in any university and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except when due reference is made in the text of the thesis.

D. D. Hutchinson

Dangnoi Dana Hutchinson

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to express my gratitude to the numerous individuals who have taken a keen interest in my work and have given me their support throughout the course of this project. I am especially indebted to my supervisor, John R. Madden, for his guidance and kind encouragement and to the staff and students of the Economics Department, who have been so generous with their comments and advice. I also wish to thank Dr. John A. Davidson for his help in my early struggles with discriminant analysis, and a special thanks goes to Graham and Wendy Fiddaman without whom I could not have typed this dissertation. Most of all, I wish to thank my husband 'Hutch' for his unending support and understanding, without whom this work would not have been possible. I dedicate this work to my parents, who had taught me from an early age that the principles of economics are not merely a field of study, but also govern everyday life.

ABSTRACT

This study makes use of two statistical methods, namely cluster analysis and discriminant analysis, to attempt to identify significant statistical relationships which may exist between various structural characteristics which are perceived to be of importance to the Tasmanian economy and certain aspects of the economic performance of manufacturing activities operating within the economy. Cluster analysis is used to group individual industries into various performance groups, after which discriminant analysis is applied to identify other attributes or characteristics, if any, that effectively distinguish between the clusters.

Data is constructed at the 3-digit Australian Standard Industry Classification level for Tasmanian manufacturing industry. The main time period under study is 1975 to 1982, although use is also made of data from earlier years when available. The data is mainly in the form of annual figures, averaged over a number of years.

The introductory chapter discusses the objectives of the study and provides a brief explanation of the methodology and approach adopted. The second chapter includes a general overview of the Tasmanian economy and reviews the main reports and studies that have considered the performance of the Tasmanian economy. The purpose of this review is not only to place the

present study in context but also to present some of the main notions concerning relationships between structure and performance variables in the Tasmanian economy which have been advanced to date. These notions formed part of the process by which variables were selected for analysis. The methodology and discussion of variables are given in Chapter 3, followed by a fourth chapter reporting the results of this study. Finally, a summary of the work done and the conclusions that can be drawn are presented in the last chapter.

CHAPTER 1. INTRODUCTION

The relationships between the structural characteristics of an economy and its performance is of interest to government bodies who are concerned with policy formulation and hence require information for decision-making. It is also of interest to private businesses -- both to those currently active in the Tasmanian economy and to those which may be potential investors in a particular activity.

A number of claims have often been made concerning how various structural features of the Tasmanian economy affect its performance. Some of the more popular of these claims are discussed in Chapter 2. However there is a lack of quantitative analysis on structure-performance links within the Tasmanian economy to provide relevant information which may substantiate or refute these claims. Apart from the Callaghan inquiry into the "Structure of Industry and the Employment Situation in Tasmania" and Wilde's work on "Industrial Structure and Change in Tasmania", no other major work appears to have been done on structure-performance links with respect to the Tasmanian economy.

It is the purpose of this study to attempt to partly fill this gap by examining some of the structure-performance relationships within the Tasmanian manufacturing sector in order to provide some insight into the nature of these relationships. This study project will attempt to identify significant relationships which may exist between various structural characteristics which are perceived to be of importance and certain aspects of economic performance, through the use of two statistical techniques, namely cluster analysis and discriminant analysis, with no presupposed model of the Tasmanian manufacturing sector. Cluster analysis is used to group the individual industries into various performance categories, and then discriminant analysis is applied to identify any variables which distinguish between these performance categories. It is hoped that the information to be gained from this study will provide a useful background for further and more comprehensive research into this area.

Cluster analysis is a method by which individual cases are grouped into a number of categories under a specified criterion. Although numerous clustering methods exist, the basic procedure is to relocate individuals to the 'closest' group determined by the specified criterion until an optimum grouping is reached, i.e. a situation where there is an optimum number of distinct groups whose members are as close as possible to the group nucleus, or centroid. In the context of this study, cluster analysis will be used to group individual industries into various performance categories, i.e. to cluster them on the basis of some dimension of performance. Having obtained these clusters of industries, discriminant analysis will be applied to identify other attributes or characteristics, if any, that effectively distinguish between the clusters.

Discriminant analysis is a method by which the group means of the observations on the selected attributes or characteristics can be tested for significant differences. It identifies the variables which have significantly different group means, from which a linear discriminant function or a set of discriminant functions can be derived to distinguish between the groups. The standardised coefficients of the function(s) show the relative contribution of each discriminating variable and as a test of the adequacy of the discriminant function(s) a set of classification equations can be derived to predict the group membership of any individual industry, given its scores on the discriminating variables. The number of classification equations is equal to the number of groups or clusters. Thus when there are four groups, an individual will have four group membership probabilities computed, and will be assigned to the group for which it has the highest probability. As a test for the adequacy of the discriminant function(s) the predicted group is compared to the individual's actual group membership. Thus a further application of discriminant analysis is to predict, for example, which performance category a new industry is likely to fall in, given its values on the discriminating variables. This latter application of discriminant analysis is beyond the scope of this study, which is limited to identifying significant statistical links between various structural characteristics and certain dimensions of performance.

Cluster analysis enabled the industries to be grouped into, in general, four distinct categories on the basis of a growth rate performance variable. The first group comprises the 'high flyers' which contains those industries which have exhibited outstandingly high growth rates in terms of value-added over the time period under study. The second group was called the 'moderate growth' industries. The third group comprised the 'stable' industries and the last group included the 'declining' industries.

Having determined these four performance categories, discriminant analysis was applied to determine if there were any statistical links between the performance groups and certain structural characteristics. The variables to be tested include those relating to energy usage, employment characteristics, protection levels and export levels. It would have been desirable to have included variables relating to transport costs, technological change, etc., but the information requirements of using such variables could not be met. Nevertheless, given the available data, certain significant results did emerge, and these will be discussed in the later sections on results and conclusions.

The original idea for this project derives from the work of T.G. Parry in his paper on "the Structure and Performance of Australian Manufacturing Industries". Parry used regression analysis as well as cluster analysis and discriminant analysis to identify links between structural characteristics and various aspects of performance. He chose 3 indicators of performance :

1) productivity

2) profitability

and 3) trade performance.

Under the framework of a small, protected international oligopoly model, he regressed the first two indicators of performance on structural characteristics such as capital intensity, natural resource inputs, 'quality of labour', concentration, etc. With respect to trade performance, he used cluster analysis to group the industries into 3 performance categories :

1) high export intensity

2) high import sensitivity

and 3) low trade involvement.

Parry then applied discriminant analysis to test for variables which distinguished between the groups.

There are 3 reasons why discriminant analysis is preferred to standard regression analysis for this present study. The first reason is that discriminant analysis does not require a formal model explaining the relationships between structure and performance. It is merely a sophisticated technique for testing for significant differences in the group means of variables, as will later become apparent. The second reason is that unlike in regression analysis, multicollinearity among the variables does not affect the results of discriminant analysis. The third reason is that discriminant analysis can be easily extended for classification and prediction purposes, although this particular aspect is not included in this study.

Data is constructed at the 3-digit ASIC level for Tasmanian manufacturing industry. This comprises 37 industry classes henceforth referred to as 'industries' for simplicity. The main time period under study is 1975 to 1982, although use is also made of data from earlier years when available. The data is mainly in the form of annual figures, averaged over a number of years.

The following Chapter gives an overview of the structure and performance of the Tasmanian economy and concludes with a brief review of studies and reports that have been conducted in this area. Chapter 3 discusses the methodology and the variables used, with a more comprehensive treatment of the design and the applications of cluster analysis and discriminant analysis. The next Chapter reports the results of this study. Finally, a summary of the work done and the conclusions that can be drawn are presented in the last Chapter.

CHAPTER 2. THE TASMANIAN MANUFACTURING SECTOR

This chapter serves to provide some background for this study of structure-performance links within the Tasmanian manufacturing sector. It is helpful to start with an overview of the Tasmanian economy, which will set the scene for the approach adopted in this particular work. Since the original idea was derived from Parry's work, this chapter will also take a closer look at Parry's paper before turning to other reports and studies which have been conducted on Tasmanian industry.

2.1 GENERAL OVERVIEW OF THE TASMANIAN ECONOMY

Tasmania is an island state which has a limited range of industries based primarily upon its natural resources. The level of unemployment is high, about 11 percent in early 1984, and in 1980 Tasmania was said to have the lowest per capita income and highest level of outmigration in Australia (Tasmanian Yearbook, 1980, p.540).

The structure of industry in Tasmania is dominated by a few large mining and manufacturing enterprises decentralised throughout the state. In general, economic activity is centered upon the state's natural resource base. In 1976, natural resource-based primary or secondary industry provided 80 percent of Tasmania's exports to the mainland and to overseas destinations (Wilde, 1981, p.221). By 1981-82 this figure had risen to about 95 percent (Tasmanian Yearbook, 1984, p.430). Wilde (1981, p.222) found that resource-based industry is specialised into only a few subsectors : 70 percent of farms rely on sheep or cattle, five minerals dominate the metallic ores sector and only four leading manufacturing industries account for 70 percent of the state factory workforce.

Tasmania has problems in attracting new industries and in enabling existing ones to maintain their commercial viability. One obstacle is that new industries would need to export a large proportion of their production because of the limited Tasmanian market. This results in what is often considered to be the root of Tasmania's economic problems : the isolation of the island and the resulting high transport costs, which puts the island industry at a disadvantage in comparison to mainland competitors. The advantages that Tasmania has to offer would therefore have to offset the transport cost disadvantage in order for new industries to establish themselves on the island.

Recent Tasmanian Yearbooks list some of the advantages that Tasmania might offer as a site for industry :

- the availability of competitively priced, bulk hydro-electricity for power-intensive industries
- abundance of natural resources, locally available raw

materials

- greater stability of the workforce than any other Australian state, in terms of industrial disputes
- availability of industrial land, harbour and shipping facilities
- abundance of water resources.

The question that remains for debate is what types of industry Tasmania should try to encourage. More knowledge about past and present structure-performance relationships could aid decision-making with respect to the development needs of the State. Once the nature of such relationships becomes apparent, more conclusive evidence can be provided to either support or refute existing claims about the Tasmanian economy. It is envisaged that more comprehensive studies along the lines of this present work could in future enable the identification of at least some industrial activity that would be beneficial to Tasmania.

At this point it would be appropriate to consider Parry's work using this particular approach and then some of the claims that have been made regarding the Tasmanian economy.

2.2 PARRY'S STUDY

Parry's (1977) report on the "Structure and Performance of Australian Manufacturing Industries" was the main inspiration for the present study. Parry used regression analysis and cluster analysis combined with discriminant analysis to identify statistical relationships between various characteristics of the Australian manufacturing sector and certain indicators of performance.

For this purpose he used data at the four-digit ASIC level for Australian manufacturing industry, and within the framework of a small, protected international oligopoly model, he chose various characteristics such as market concentration, factor intensity, diversification, etc. to test for variables which influence performance. It should be noted here that discriminant analysis is not a suitable method for identifying causal relationships, but rather for merely identifying statistically significant relationships. The three indicators of performance which he used were productivity, profitability and trade performance, for which only trade performance was used in cluster analysis. In other words, Parry used standard regression analysis to test for behavioural relationships between various structural characteristics and profitability and productivity within the framework of his model, and then switched to a different approach in which he applied cluster analysis to group the industries into performance groups on the basis of trade performance. Having done that, he then went on to apply discriminant analysis to the clustered industries in search of characteristics that would distinguish between the groups.

From his regression analysis, Parry found that the human resource aspect was a key factor in productivity, in that labour skills were a significant and positive factor, while migrant and female worker utilisation rates were significant and negative factors. Other significant and positive relationships to productivity were found with capital inputs, natural resource inputs, concentration, number of small enterprises, diversification and multi-plant operations. From his regression analysis on profitability, Parry found that diversification, change in average rates of protection and foreign ownership were inversely and significantly related to productivity.

With regard to trade performance, Parry used cluster analysis to group industries into three categories : 1) high export intensity 2) high import sensitivity and 3) low trade involvement. Using discriminant analysis he found that the highly import sensitive group was characterised by high migrant and female employment, high average wages of administrative workers, high research and development expenditure and by foreign ownership. Natural resource use, effective protection and oligopoly-model characteristics were good discriminators in sorting industries to the low trade involvement group. From his results Parry also concluded that industries associated with high effective rates of protection are characterised by their capital intensity, preponderance of small firms with high migrant and female participation rates, capital-city concentration and outward diversification, and that the role of tariff protection could be regarded as an effective policy instrument in protecting the domestic market from imports, with consequent effects on export performance.

The present study follows Parry's approach in using cluster analysis and discriminant analysis to identify significant statistical links between structural characteristics and performance indicators of Tasmanian manufacturing industries.

2.3 CLAIMS CONCERNING STRUCTURE-PERFORMANCE LINKS WITHIN THE TASMANIAN ECONOMY

Probably the best report concerning the Tasmanian manufacturing sector is Callaghan's 1976 "Inquiry into the Structure of Industry and the Employment Situation in Tasmania". One of the conclusions of the inquiry was that "...there is little scope for significant expansion in the manufacturing sector as things stand at present" (p.34). The inquiry also found that previous Tasmanian government policies to encourage the establishment of power-intensive industries had approached its limits in terms of the availability of water resources for hydro-electricity, and that other traditional sources of power, such as oil, gas and coal were also limited. Another significant finding was that "the structure of manufacturing industry differs significantly in Tasmania from the mainland in terms of size, location, marketing and specialisation" (p.45). Thus the largest Tasmanian industries are more closely related to natural resources than the largest mainland industries, and the Tasmanian ones tend to be more decentralised. This, according to Callaghan, has resulted in the development of regional communities around the larger factories and mines, and hence vulnerability of the Tasmanian economy to a few large industries. In short, Callaghan felt that the tertiary sector had the most potential for growth and development, and that until further employment opportunities arise in that sector, large scale manufacturers should be encouraged.

The only other major work that has been done specifically in the field of industrial structure and performance is that of Wilde (1981) who views the Tasmanian economy in the context of a core-periphery model, where the core region comprises a diverse and sophisticated industrial structure while the peripheral regions are "...resource-based, typically specialised on a single resource subsector, and dependent for survival and growth on their ability to supply raw or minimally processed natural resources to the core at competitive prices" (p. 219). Wilde argues that the natural resource-based industries have resulted in an intensely specialised industrial structure in Tasmania, and in "...failure to participate proportionally in growth generated in managerial, manufacturing or natural resource sectors" (p. 220). He argues that a share of manufacturing industry is 'filtered down' from the core, i.e. from the mainland regions of New South Wales, Victoria and the A.C.T., to peripheral regions such as Tasmania, so that the growth potential of these 'filtered down' industries have been virtually exhausted.

Wilde uses a 'shift and share' analysis to compare Tasmanian trends with national trends in individual subsectors. The technique distributes total employment changes in each sector of the Tasmanian economy between a 'regional share' (i.e. what would occur if the industry had expanded at the national growth rate), and 'industry performance effect' (i.e. the extra increase or decrease in the industry's employment if it had expanded at the national rate), and a 'competitive effect' (i.e. the difference in actual Tasmanian performance for each industry from the expected change based on national growth rates). He found that rural, mining and manufacturing industries were subject to erratic growth and had poor growth prospects in terms of national performance, and that employment was also characterised by poor prospects. His findings also agreed with Callaghan's in that it was the tertiary sector that appeared to have the greatest potential for development.

The recent debate over the further development of water resources for hydro-electricity generation has also raised some questions concerning the future structure of industrial activity in Tasmania. The Hydro-Electric Commission is of the opinion that increased generation capacity is needed to meet the future industrial requirements of the State. Indeed, as stated earlier in this chapter, abundant and competitively-priced hydro-electricity has been traditionally regarded as an incentive for industries to site in Tasmania. However the competitiveness of price relative to other States appears to be in some doubt (see Jones, 1980) and the employment aspect of attracting large, power-intensive industries is also questioned. Apart from Callaghan's and Wilde's pessimistic predictions of slow overall growth and employment growth in Tasmanian manufacturing, the 1982 Tasmanian Yearbook also foresees a gloomy future for high energy consuming industries. Furthermore, it points out an apparent conflict of government objectives, in that over the past decade sales of electricity to industrial consumers rose by 51.6 percent while employment in manufacturing has declined by 17.4 percent :

> "Employment decline has been particularly marked in the high energy using industries of basic chemicals, chemicals and paper production. Tasmania faces the problem of a small local market and transport difficulties in its attempts to attract and retain industry. These factors have encouraged the development of those industries able to take full advantage

of cheap power, such undertakings, however are not necessarily large employers of labour." (p.235)

There is little statistical evidence available to back claims relating to the performance of the Tasmanian economy. Thus the motivation for this research is to provide some statistical analysis to test at least some of the assertions which relate to links between structure and performance.

CHAPTER 3. METHODOLOGY AND DISCUSSION OF VARIABLES

In this study it has been assumed that there is no prior knowledge of which characteristics may be linked with which levels of performance. It is the purpose of this study to identify the statistical relationships which may emerge as being significant in structure-performance links. It is not the purpose of this study to suggest policy recommendations, but rather to provide useful information for further research on the nature of structure-performance relationships.

The approach adopted consists of two main stages : first, cluster analysis to group Tasmanian manufacturing industries into various performance categories, and second, discriminant analysis to identify any characteristics which may effectively discriminate between the different categories of industry, and also how well they do discriminate.

Obviously the success of such analysis will depend, inter alia, upon the completeness and availability of data and on the suitability of the variables chosen for analysis. The data used consists of cross-section observations on Tasmanian manufacturing industries, mainly over the period 1975 to 1982, but experiments were also conducted for the period 1968 to 1974.

The variables which were tested were selected firstly on the basis of their suitability for testing relationships between structure and performance that might bear on such propositions as referred to in the previous chapter, plus a number of other plausible relationships, but unfortunately also on the second criterion of the availability of data. Wherever possible the observations were averaged over a period of time in order to be as representative as possible of each industry. It would have been desirable to have used data at the 4-digit ASIC level of disaggregation, but due to the size of the Tasmanian manufacturing sector, 3-digit level data had to be used. Even at this level there were problems with missing observations due to confidentiality requirements. Wherever possible, these missing values were estimated, although in certain circumstances which will be discussed below, missing data may not seriously affect the outcome of discriminant analysis.

First, the technique of cluster analysis will be discussed, and then discriminant analysis. This will be followed by a discussion of the selected variables.

3.1 CLUSTER ANALYSIS

Cluster analysis is described by Wishart (1978, p.1) as "...an exploratory method for helping to solve problems" with the objective of sorting "...a sample of cases under consideration...into groups such that the degree of association is high between members of the same group and low between members of different groups". It is widely used in zoology and botany for classification purposes, and has more recently been applied in social sciences. In the context of this study it means that we attempt to group industries of similar performance levels together and hope to discover, via discriminant analysis, if there are certain structural characteristics which distinguish between these groups. Thus with the information obtained, it may be possible to form an opinion of the possible performance of an industry given certain characteristics.

There are several different clustering techniques which are widely used. Parry used a non-hierarchical cluster analysis, which is an iterative technique in which observations continue to be relocated amongst the designated number of groups until the selected criterion has been optimised. The particular criterion used by Parry was the minimisation of the determinant of the pooled within-group deviation sum of squares, a Euclidean measure of distance which results in a multivariate clustering algorithm for spherical clusters.

This study follows Parry's method of non-hierarchical clustering with the criterion of the minimisation of the pooled within-group sum of squares. The basic procedure is to choose an initial partition of the data units, in this case, industries, and then relocate individuals to different clusters in order to obtain a better partitioning. Anderberg (1973, p. 156) compares this procedure to algorithms used for unconstrained optimisation in non-linear programming, i.e. "...such algorithms begin with an initial point and then generate a sequence of moves from one point to another, each giving an improved value of the objective function, until a local optimum is found".

In cluster analysis, the 'objective function' is the selection criterion. The error sum of squares is given by

$$E_k = \sum_{j=1}^{m} \sum_{i=1}^{n} (x_{ijk} - \bar{x}_{ijk})^2$$

where E_k = error sum of squares of cluster (k), also called the "Euclidean" or "within-group" sum of squares, i.e. it is the sum of squared deviations about the centroid of cluster (k)

x_{ijk} = score on the (i)th of (n) variables for the (j)th of (m) cases in the (k)th of (h) clusters

$\bar{x}_{ijk}$ = $\sum_{j=1}^{m} x_{ijk}/m$ i.e. mean of the (i)th variable for the member cases of cluster (k)

Thus the total within-group error sum of squares for the (h) clusters is given by

$$E = \sum_{k=1}^{h} E_k$$

According to Wishart (1978, p. 115), this method "...is suitable for finding tight clusters which have the property that each cluster represents the constituent cases at a high level of similarity with respect to all the underlying variables".

Individual industry performance was assessed in terms of "relevant growth rates". The "relevant growth rate" of individual manufacturing industries is approximated by the rate of change in the proportion of value-added of that industry to the total value-added of the manufacturing sector.

Thus the 32 industries which were included in the analysis were randomly assigned to 10 initial clusters, or groups. Next they were relocated one at a time, under the minimisation of the pooled within-group sum of squares criterion, to different clusters, with the most similar clusters fused together to form new clusters. This method is called iterative relocation with hierarchic fusion, and is repeated until the specified number of terminal clusters is reached, in this case, 3.

The advantage of using a clustering method such as this is that it enables us to categorize individuals on a systematic basis using a specified and precise mathematical criterion, a process which can become very complex when the individuals are to be grouped on the basis of more than one variable.

Once this has been done, we can then turn our attention to the characteristics of each group. Discriminant analysis was then used to provide information on two main questions :

1) are various characteristics of the clustered groups significantly different ? and

2) if they are, which of these characteristics contribute most to the difference?

3.2 DISCRIMINANT ANALYSIS

Discriminant analysis is a statistical method which is used to assign cases to one or more populations on the basis of a discriminant function. A discriminant function is a linear combination of discriminating variables, i.e. some set of characteristics with coefficients estimated from sample data.

Discriminant analysis can be used for two main purposes, interpretation and classification. The first of these, interpretation, involves an examination of the nature of group differences between two or more groups, in terms of a set of characteristics, in order to determine how well each of the characteristics discriminate between groups both individually and grouped with other characteristics. The second purpose of discriminant analysis, classification, makes use of discriminant functions or of discriminating variables in such a way as to identify a group which a case most closely resembles. Classification can also be used as an indicator of the performance of the discriminant functions and the use of classification in this study will be limited to that.

The mathematical form of a discriminant function is given by

$$d_{km} = u_0 + u_1 Z_{1km} + \dots + u_p Z_{pkm} \qquad [1]$$

where d_{km} = score on canonical discriminant function for case (m) in group (k)

Z_{ikm} = value of discriminating variable (i) for case (m) in group (k)

u_i = coefficient which produces desired characteristics

u_0 = constant

we also define

n = total number of cases

n_k = number of cases in group (k)

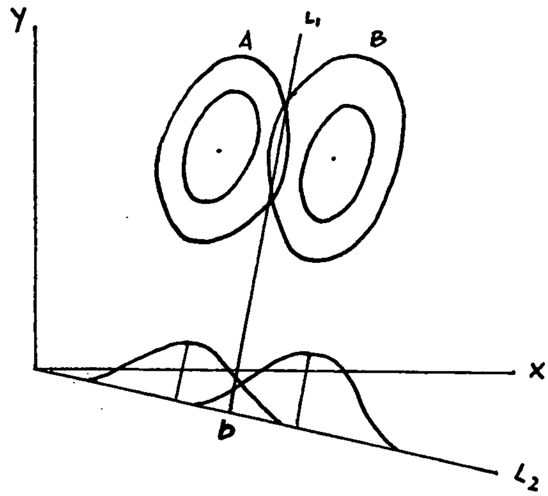
g = total number of groups

p = total number of discriminating variables

The objective is to obtain a linear combination of variables that will optimally classify observations into one group or another. This is not the same as grouping observations under cluster analysis. For example, given a sample of industries, we can cluster them on the basis of a particular attribute, say, performance. If we obtain only two groups (e.g. high and low performance) we can then use discriminant analysis to find out if there are other attributes (i.e. characteristics) which determine to some extent to which performance group any particular industry belongs.

For each discriminating variable (i) the mean value over all cases in the sample is referred to as the grand mean of that discriminating variable and is denoted by $\bar{Z}_i$. When these grand means are substituted into the discriminant function they determine the critical value of "d", i.e. d*. For any new industry, the values of its characteristics (i.e. discriminating variables) are multiplied by their corresponding coefficients taken from the discriminant function, and the sum of these products yield the score for that industry. The industry's score is then compared to the critical value of d (i.e. d*) and if it is less than d* it is assigned to one group, and if it is greater than d* it is assigned to the other group. Where there are more than two groups, the individual is assigned to the group whose score (calculated by inserting the group means into the discriminant functions) is closest to the individual's score.

Cooley and Lohnes (1971, p.245) set out an exceedingly simple and clear graphical presentation of the nature of discriminant analysis, the essence of which is given below.



Suppose we have two populations, A and B, and that for each member of the populations we have observations on two variables, X and Y. Let the 2 sets of concentric ellipses represent the bivariate distributions for the 2 populations A and B in standardised form, such that the outer ellipses define the regions where (say) 90 percent of each population lies while the inner ellipse defines the region where (say) 75 percent of each population lies. The 2 points at which the 2 outer ellipses intersect define the straight line L_1. If we draw a second line L_2 perpendicular to L_1 and project the points in the two-dimensional space onto L_2, we obtain the smallest possible overlap between the two populations. L_2 is in fact the discriminant function which transforms individual scores into a single discriminant score, and that score is the individual's location along L_2. Point **b** divides the one-dimensional discriminant space into 2 regions, one indicating probable membership in population A and the other for population B.

Bolch and Huang (1974, p. 231) give a three-dimensional treatment of discriminant analysis, and interested readers are referred to Appendix B.

The key element in the interpretation of discriminant functions is the set of coefficients. In brief, it is through the definition and derivation of these coefficients that the procedures for determining the usefulness of the discriminant functions are obtained. There are four indicators which aid in the interpretation of discriminant functions : standardised coefficients, eigenvalues and relative percentages, canonical correlation coefficients and Wilks's Lambda (or the U-statistic). The first is used to identify the most powerful discriminators while the remaining three relate to the degree of discriminating power of the discriminant functions.

It has been shown above that the derivation of discriminant functions is actually a transformation from a p-dimensional space [i.e. (p) discriminating variables] to a q-dimensional space [i.e. (q) discriminant functions]. We should note that (q) denotes the maximum number of unique functions that can be obtained and is equal to either (g-1) or (p), whichever is the smaller. Thus the score on each discriminant function for any given case represents its coordinates in the space defined by that function. Sometimes a discriminant function will define a space which overlaps with that already defined by another function. In other words, that function is redundant because it does not provide any new information. However, the overlap is often not complete due to sampling and measurement errors. Therefore each function should be tested for statistical significance, and the means for testing, i.e. Wilks's Lambda (sometimes referred to as the U-statistic), is obtained through the derivation of the discriminant function coefficients.

The coefficients of discriminant functions are derived so that they provide a measure of the degree of differences among data cases. This entails measuring within-group dispersion and between-groups dispersion. The underlying principle is that when group locations are indeed distinct the degree of dispersion within groups will be less than total dispersion. Thus we can define two square symmetric matrices **T**, representing total dispersion and **W**, representing within-group dispersion such that

$$t_{ij} = \sum_{k=1}^{g} \sum_{m=1}^{n_k} (\overline{Z}_{ikm} - \overline{Z}_i)(\overline{Z}_{jkm} - \overline{Z}_j) \qquad [2]$$

where $\overline{Z}_i$ = mean value of discriminating variable (i) for all cases, i.e. "grand mean" for variable (i)

and $$w_{ij} = \sum_{k=1}^{g} \sum_{m=1}^{n_k} (\overline{Z}_{ikm} - \overline{Z}_i)(\overline{Z}_{jkm} - \overline{Z}_j) \qquad [3]$$

where $\overline{Z}_{ik}$ = mean value of discriminating variable (i) for all cases in group (k) only

We can further define a matrix **B** = **T-W** so that the size of **B** relative to **W** yields a measure of how distinct groups are. We can denote this relationship as a set of simultaneous equations in the form of

$$\begin{aligned} \Sigma\, b_{1i} v_i &= \lambda\, \Sigma\, w_{1i} v_i \\ &\vdots \\ \Sigma\, b_{pi} v_i &= \lambda\, \Sigma\, w_{pi} v_i \end{aligned} \qquad [4]$$

where λ = constant called an eigenvalue

v_i = set of (p) raw coefficients

Therefore each set of λ and (v)s corresponds to one discriminant function. When the situation arises that group centroids are identical, i.e. when a function defines a space completely overlapping that already defined by another function, **B** will be equal to zero because within-group dispersion will be the same as total dispersion, i.e. **B** = **T-W** = 0. Therefore the eigenvalue will also be zero. Hence when the eigenvalue is close to zero we are faced with the task of testing for the statistical significance of the perceived differences. In other words, we need to find the probability of having obtained cases which show the computed degree of discrimination when in fact no real differences exist.

Rather than test the function directly we can examine the degree of residual discrimination, i.e. the "ability of variables to discriminate between groups after information has been extracted by previously computed functions" (Klecka, 1980, p.38). Residual discrimination is measured by Wilks's Lambda, or the U-statistic, which can best be visualised as a multivariate F-ratio, in the sense that it compares the differences between the vector of group means on the discriminating variables of one group with that of another. The conventional formula for Wilks's Lambda is given by

$$\mathbf{U} = \prod_{i=k+1}^{q} \frac{1}{1 + \boldsymbol{\lambda}_i} \qquad [5]$$

where k = number of functions already derived

$\prod$ means that individual terms are multiplied to yield the final product

Thus **U** is an inverse measure, such that as **U** approaches zero it indicates high discriminating power amongst the variables because the eigenvalues are large, reflecting distinct groups. The converse is true as **U** approaches unity.

Wilks's Lambda, or **U**, can also be converted into a test of significance in the form of either an F or a chi-square distribution, so that standard tables can be used to determine significance levels. Conventionally, the chi-square distribution is used because of convenience in calculation. A formula which relates $\mathbf{U}_k$ to chi-square is

$$\mathbf{X}^2 = -[n - \left(\frac{p+g}{2}\right) - 1] \ln \mathbf{U}_k \qquad [6]$$

with (p-k)(g-k-1) degrees of freedom.

The null hypothesis is that there is no difference in the vectors of group means of the discriminating variables and the alternative hypothesis is that there is a difference. The computed value of chi-square is then compared to the critical value given in standard tables at certain levels of significance, and if the computed value equals or exceeds the critical value the null hypothesis is rejected. Thus, given a level of significance of, say, .01, we can reject the null hypothesis with a one percent chance that the decision to reject it is wrong, i.e. we can assume that the perceived differences are real and significant.

The eigenvalues obtained from the computation of discriminant function coefficients can also be used to derive two other measures of the discriminating power contained by the functions. The first measure is called the relative percentage of each function. This is obtained by dividing the eigenvalue of each function by the sum of all the eigenvalues of all functions derived so that the final figure represents a measure of the total discriminating power of that function in the system. Thus a relative percentage of, say, .90 means that the particular function in question accounts for 90 percent of the total discriminating power of all the functions derived and is therefore by far the most important function.

The second measure of discriminating power is the canonical correlation coefficient which measures the degree of relatedness between groups and discriminant function (Klecka, 1980, p.36). The canonical correlation coefficient, r^*_i, is given by

$$r^*_i = \sqrt{\frac{\lambda_i}{1+\lambda_i}} \qquad 0 \leq r^*_i \leq 1 \qquad [7]$$

Recall that when λ is close to zero the groups are not very distinct so that the functions define overlapping spaces. Therefore as r^* approaches zero we can say that there is probably no relationship, i.e. that the function does not really discriminate between the groups. Conversely, as r^* approaches unity we can be more certain that the function does indeed discriminate between the groups.

Having identified the functions which do discriminate between groups, thereby establishing that characteristics do distinguish between groups, we can proceed to examine the discriminating variables included in the functions in order to identify which variables are the most powerful discriminators in terms of highest contribution to separation of the groups. This is indicated immediately by the coefficients of the discriminant functions. When raw data is used in the derivation of discriminant functions we obtain unstandardised coefficients. They are 'unstandardised' because raw, i.e. unstandardised, data was used in the computations.

The unstandardised coefficients, denoted by u_i, are obtained from the set of (p) raw coefficients, denoted by v_i) from Equation [4], by the following transformation :

$$u_i = v_i \sqrt{n-g} \qquad \text{and} \qquad u_0 = -\sum_{i=1}^{p} u_i \bar{Z}_i \qquad [8]$$

This results in unstandardised coefficients with the property that the resulting discriminant scores for the data cases are in standard form, i.e. the scores over all cases will have a mean of zero and a within-group standard deviation of unity.

Unstandardised coefficients show the absolute contribution of each variable, but if each variable had been measured in different units (as often the case would be) it is impossible to determine which of the discriminating variables are the most important. More meaningful comparisons can be made if the data were standardised, for we would then obtain standardised coefficients which show the relative contribution of each variable. In order to obtain standardised coefficients we could either standardise the raw data or alternatively we could use the following transformation to convert unstandardised coefficients into standardised coefficients :

$$s_i = u_i \sqrt{\frac{w_{ii}}{n-g}} \qquad [9]$$

where u_i = unstandardised coefficients

s_i = standardised coefficients

w_{ii} = corresponding element in the matrix for within-group differences, **W**

3.2.1 CLASSIFICATION

We can then examine the absolute value of each standardised coefficient to determine the magnitude of its contribution. The greater the absolute value of the coefficient, the greater its contribution to the separation of the groups.

The classification stage of discriminant analysis can now be used to test the accuracy of the discriminant functions. This can be done either with the discriminant functions themselves or with the discriminating variables. The most convenient method is to use the discriminating variables to obtain a separate classification equation for each group. The classification equations are of the form

$$C_k = c_{k0} + c_{k1}Z_1 + \ldots + c_{kp}Z_p \qquad [10]$$

where C_k = classification score for group(k)

c_{k0} = constant

c_{ki} = classification coefficients derived from the elements of the inverse of the within-group matrix **W** and the means of the discriminating variables

Z_i = raw score on discriminating variable (i)

The classification functions can then be applied to individual cases in order to determine to which group a case is most likely to belong.

Thus for each individual case, a classification score will be computed for each group, and the individual will be assigned to the group for which it has the highest score. Under the assumption of a multivariate normal distribution, i.e. when the observations on the variables have been standardised, the classification scores are converted into probabilities of membership for each group. Therefore the individual case will be assigned to the group for which it has the highest membership probability.

It should be noted that the use of a set of data to compute a discriminant function and the application of that function to the same set of data results in an upward bias in the function (see Eisenbeis, 1977 and Conlon, 1983). The bias is due primarily to sampling errors in estimating the means of the population, and will result in greater predictive power in classification than would actually exist given the true populations. Therefore when the sample size is small there would tend to be a greater amount of bias, and this would lessen as the sample size is increased. It is because of this that missing values for some individuals do not present a serious problem in the analysis. For example, Conlon (1983) randomly selected 86 out of 170 industries at a time upon which to compute discriminant functions and used the resulting functions to classify the remaining 84 industries. On the one hand, such a procedure enables the functions to perform more accurately at the classification stage, but on the other hand, when the sample size is small to start with, as in the case for this study, we cannot benefit from using a larger sample to compute the functions.

3.2.2 STEPWISE DISCRIMINANT ANALYSIS

The previous section has provided a brief introduction to the nature of discriminant analysis by examining the discriminating power contained in individual discriminating variables. Discriminant analysis can also use various combinations of variables to derive discriminant functions. The underlying principle is that while individual characteristics may not discriminate adequately between groups, a combination of variables may do so. For example, if one wanted to identify the characteristics which would distinguish between Caucasians and Asians, hair colour, skin colour or height by themselves may not be sufficient to determine to which group an individual would belong. However, all three characteristics taken into consideration simultaneously may provide a better 'discriminant function'. Hence we can attempt to derive multivariate discriminant functions. The question then becomes which characteristics should be included in these functions? Stepwise discriminant analysis provides a means for selecting the 'best' discriminating variables to be included in the discriminant functions.

In a stepwise procedure independent, or discriminating variables are entered on the basis of their discriminating power. As Klecka (1975) explains:

> "The process begins by choosing the single variable which has the highest value on the selection criterion. This initial variable is then paired with each of the other available variables, one at a time, and the selection criterion is computed. The new variable which in conjunction with the initial variable produces the best criterion value is selected as the second variable 'to enter the equation'. These two are then combined with each of the remaining variables, one at a time, to form triplets which are evaluated on the criterion. The triplet with the best criterion value determines the third variable to be selected. This procedure of locating the next variable that would yield the best criterion score, given the variables already selected, continues until all variables are selected or no additional variables provide a minimum level of improvement." (p. 447)

Prior to testing by the selection criterion, which will be discussed below, a variable is generally required to pass two minimum conditions : 1) a tolerance test to ensure computational accuracy and 2) a partial F-statistic test to ensure that the increased discrimination resulting from the inclusion of that variable exceeds some level determined by the researcher, depending upon the situation. The tolerance test of a variable not yet selected is to subtract the squared multiple correlation between that variable and all other variables already selected (obtained from a within-group correlation matrix similar to **W** discussed above) from unity. Tolerance levels close to zero make further computation difficult and also indicate that the variable is a linear combination of the variables already selected, thus containing no new information. The partial F-statistic test actually comprises two tests : an F-to-enter and an F-to-remove. The F-to-enter is a partial multivariate F-statistic which tests the statistical significance of the additional discrimination introduced by the variable being considered, taking into account the discrimination already achieved by the variables already selected. F-to-remove tests the significance of a decrease in total discriminating power should the variable in question be removed from the list of variables already selected. In other words, F-to-remove ensures that variables that have become redundant are removed from the list of selected variables because their contribution to discrimination has since been duplicated by other variables.

Once variables have satisfied the minimum conditions for selection they are then tested on the selection criterion. There are five conventional criteria : Wilks's Lambda and the partial F-statistic, Rao's V, Mahalanobis's D, between-groups F and minimising residual variance.

The first of the selection criteria, Wilks's Lambda or the U-statistic takes into consideration both differences between groups and homogeneity within groups. (Recall Equation [5]). Since **U** is an inverse measure, the variable with the smallest value for **U** will be selected. Wilks's Lambda can also be converted into an F-statistic to test for group differences, so that the variable selected will be the one with the largest value of F. Alternatively, the partial F-statistic, or F-to-enter as discussed above, can be used. All three statistics yield the same result.

The second possible criterion is Rao's V, a generalised measure of distance. It measures the separation of group centroids and is given by

$$V = (n-g)\sum_{i=1}^{p'} \sum_{j=1}^{p'} a_{ij} \sum_{k=1}^{g} nk(\bar{Z}_{ik} - \bar{Z}_i)(\bar{Z}_{jk} - \bar{Z}_j) \qquad [11]$$

where V = value for Rao's V

n = total number of cases

g = total number of groups

p' = number of variables already selected

a_{ij} = element of the inverse of matrix **W**, i.e. $\mathbf{W}^{-1}$

n_k = number of cases in group (k)

$\bar{Z}_{ik}$ = mean value of discriminating variable (i) for all cases in group (k) only

$\bar{Z}_i$ = mean value of discriminating variable (i) for all cases, i.e. "grand mean" for variable (i)

When the number of cases is large **V** has a sampling distribution approximately the same as chi-square with p'(g-1) degrees of freedom. In addition, the change in **V** due to the addition or deletion of variables also has a chi-square distribution with degrees of freedom equal to (g-1) times the number of variables added or deleted at that step. Therefore we can test for the statistical significance of the change in overall separation; if the change in **V** is not statistically significant it would not be desirable to include that variable.

Unlike the first two criteria which concentrate on maximising group separation, the remaining three criteria select the variable which generates the greatest separation for the pair of groups which are closest at that step. Often, however, all five criteria yield the same result, and the decision to apply the Rao's V criterion in this study was based on the grounds that it provided the clearest rationale for the identification of characteristics which would distinguish between the four industry groups.

As the results of discriminant analysis may become difficult to interpret when there are more than two groups or clusters involved, the approach taken is to undertake discriminant analysis firstly for all four performance groups, and then for all pairs of groups. Answers are sought to questions such as what combination(s) of characteristics distinguish the high flying industries from the declining industries, or the declining industries from the stable ones. Stepwise discriminant analysis is the technique applied in an attempt to answer some of these questions.

3.3 DISCUSSION OF VARIABLES

The structure of an economy refers to the characteristics of the environment which influence the behaviour of the participants and covers such things as the level of protection, the number of buyers and sellers, the type of inputs required, etc. Performance, on the other hand, is the evaluation of the resulting allocation of resources and therefore entails assessment on many aspects. It is impossible to cover all the conceivable aspects of industry performance and structural characteristics in a study of this size, but it is hoped that further research by others will be carried out in future. Hence the scope of this study is confined to the evaluation of performance only in terms of one performance indicator, growth as reflected in the rate of change in an industry's contribution to total manufacturing industry's value-added. The structural characteristics will be confined to those which appear to be most relevant to Tasmania, as suggested in the literature.

3.3.1 PERFORMANCE INDICATORS

Industry performance is a concept that embodies a host of aspects such as efficiency, profitability and growth. The evaluation of an industry's performance involves the assessment of how well an industry can achieve its objective, or the extent to which it is able to realise its potential to generate direct and indirect benefits to the surrounding economic environment.

Caves, Ward, Williams and Wright (1981) describe one view of performance in terms of a general framework of industrial organisation. This framework is based upon the relationships between the three concepts of market structure, market conduct and market performance. Market structure refers to the features of a market environment which influence the behaviour of buyers and sellers. Market conduct refers to the policies of participants towards the market in terms of price, product characteristics, etc. market performance refers to the normative appraisal of the social quality of the allocation of resources which results from a market's conduct. In assessing performance, an attempt is made to evaluate the achievement of four basic objectives : efficiency in terms of resource/factor utilisation, progressiveness in terms of development and innovation, equity in terms of income distribution and stability of prices and employment (Caves et al, 1981, p.81). They suggest that by identifying "...reliable links between elements of structure and elements of performance, we have a powerful tool for economic analysis and public policy" (Caves et al, 1981, p.11).

The lack of a suitable data base for the Tasmanian manufacturing sector necessitated the search for appropriate indicators of performance for which sufficient data was available. Growth per se may not necessarily be a suitable performance indicator for an individual industry if it is achieved at the cost of misallocation of resources through price distortion and the consequent decline of other industries. Nevertheless the concept of industry "growth" in value-added seemed to be appropriate as it is policy-oriented and relevant to the claims advanced in the previous chapter. It was originally intended to add other performance indicators such as profitability, but lack of time precluded the construction of the required data.

Hence individual industry performance was assessed in terms of "relevant growth rates". The "relevant growth rate" of individual manufacturing industries was approximated by the rate of change in the proportion of value-added of that industry to the total value-added of the manufacturing sector. Value-added is defined by the Australian Bureau of Statistics as turnover, plus increase (or less decrease) in the value of stocks, less purchases, transfers in and selected expenses, i.e. it is the basic measure of an industry's contribution to total production.

The basic procedure adopted was simply to calculate the proportion of value-added contributed by each industry class and compare it to the corresponding proportion in the previous year. Consider, for example, Tables 3.3.1A and 3.3.1B which show the relevant items for wood, wood products and furniture, and for transport equipment for the years 1974-75 and 1975-76 respectively.

Thus for the wood and wood products industry, the "relevant growth rate" for 1975-6 is given by (14.6 - 15.9)/15.9 = -8.18 %. Similarly, the "relevant growth rate" for the furniture industry is given by (1.2 - 1.1)/1.1 = 9.09 %.

The "relevant growth rate" (henceforth simply referred to as "growth rate") for each 3-digit ASIC level industry was thus calculated for the years 1975 to 1977 and for 1980 to 1982. The percentage change for each industry for each year was then averaged to obtain a final figure representing the average annual growth rate for each industry over the time period 1975 to 1982. In doing this it should be noted that only nominal values were used. It is assumed that over the relevant time period relative prices remained constant so that nominal differences in value-added reflected real changes rather than changes due to changing relative prices. Insufficient data on price movements meant that it was not possible to employ any alternative assumptions.

TABLE 3.3.1A 1974-75

ASIC	INDUSTRY DESCRIPTION	(a)	(b) ($ '000)	(c) (%)
251	Wood & Wood Products	174	64 150	15.9
252	Furniture	28	4 520	1.1
25	Wood, Wood Products & Furniture	202	68 670	
321	Motor Vehicles & Parts	12	n.p.	-
322	Other Transport Equipment	12	n.p.	-
32	Transport Equipment	24	12 648	
	TOTAL MANUFACTURING		402 255	

TABLE 3.3.1B 1975-76

ASIC	INDUSTRY DESCRIPTION	(a)	(b) ($ '000)	(c) (%)
251	Wood & Wood Products	185	66 590	14.6
252	Furniture	34	5 657	1.2
25	Wood, Wood Products & Furniture	219	72 247	
321	Motor Vehicles & Parts	12	n.p.	-
322	Other Transport Equipment	15	n.p.	-
32	Transport Equipment	27	15 082	
	TOTAL MANUFACTURING		456 029	

Source : Economic Censuses : Manufacturing Establishments Australian Bureau of Statistics, 1974-75, 1975-76.

(a) number of establishments
(b) value-added in $ '000
(c) proportion of total manufacturing value-added
n.p. not published

Generation of missing values for value added

A problem arose in calculating growth rates for industries for which only the aggregate 2-digit ASIC level data is published, e.g. in transport equipment. One method that could be used to estimate the necessary missing values is to assume that all establishments within an industry contribute equally to production, and hence take a straight proportion using the number of establishments in each industry. For example, in 1974-75 there was a total of 24 establishments engaged in producing transport equipment, with total value-added amounting to \$12, 468 (x 10^3). Since each 3-digit level industry comprised 12 establishments we can apportion half of the total value to each industry, i.e. \$6, 234 (x 10^3). However, this would be an extremely crude method.

Fortunately, additional data was available in the form of ratios reflecting the contribution of each Tasmanian industry to Gross State Product, i.e. the proportion of each industry's production over Gross State Product (henceforth referred to as GSP ratios). However, GSP ratios were only available for the base year 1974 so that in order to try to capture any possible changes in subsequent years they were modified by using the change in the number of establishments for each year after 1974. Thus a similar apportioning method to the one described above could be used to estimate the missing values, based upon the number of establishments operating in each industry, and drawing upon the additional information that was available.

Suppose in year (0) we have an industry group denoted by X^0, and that the industry group consists of (n) industries. Let x^0_i denote the number of establishments comprising industry (i) for i = 1...n in year (0), and r^0_i denote the GSP ratio for industry (i) for that year; in this case it is industry (i)'s contribution to Tasmania's Gross State Product in year (0) = 1974. Hence we know the values of

$$\text{year (0)} : x^0_i,\ r^0_i \qquad i = 1\ldots n$$

$$\text{and for any year (t)} : x^t_i \qquad i = 1\ldots n$$

We can now obtain a new figure, r^t_i, by weighting r^0_i by the change in x_i

$$\text{thus } r^t_i = \frac{x^t_i}{x^0_i} (r^0_i) \qquad [12]$$

The GSP ratios are then applied to the 2-digit level industry's value-added figure, in order to estimate the missing values at the 3-digit level, by using the formula:

for industry (i),

$$\frac{r^t_i \text{ x 2-digit industry's value-added}}{\sum_{i=1}^{n} r^t_i} \qquad [13]$$

3.3.2 SELECTED STRUCTURAL CHARACTERISTICS

The structural characteristics selected for testing via discriminant analysis were chosen because of their relevance to the claims and propositions discussed in Chapter 2 and also on the basis of their availability. It is regretted that more variables could not have been tested due to the lack of adequate data on the Tasmanian manufacturing sector, particularly at the 3-digit and 4-digit ASIC levels.

The selected structural characteristics fall into four main categories : resource utilisation, protection, export intensity and 'local element', i.e. Tasmanian-based versus mainland-based.

Resource Utilisation

The main resource utilisation characteristic centers upon the debate about the need for increased energy generation through the further development of the State's water resources and the argument that Tasmania needs to attract large power-intensive industries. Therefore the more conventional type of factor-intensity measures were rejected with respect to energy utilisation because it was felt that it is not the energy-intensiveness per se which is of concern, but rather the absolute level of energy requirements. Hence the original three energy usage variables below were intended to reflect this concern.

1) ENERG1 = ratio of the value of electricity plus other fuels for industry (x) to the total value of electricity plus other fuels for all manufacturing

2) ENERG2 = ratio of the value of only electricity for industry (x) to the total value of only electricity for all manufacturing

3) ENERG3 = ratio of the value of only electricity for industry (x) to the value of electricity plus other fuels for industry (x)

Thus ENERG1 and ENERG2 refer to the proportion of an individual industry's consumption as compared to that for all manufacturing, while ENERG3 refers to the amount of electricity an individual industry consumes relative to its other fuel sources, if any.

The experimental nature of the approach adopted in this project, in the search for characteristics that will discriminate between industry performance groups, led to the trial of a modified set of energy consumption variables. It was thought that the rate of change in energy consumption could perhaps be an important structural characteristic by reflecting the trends in energy consumption. Thus we have the modified set of energy utilisation variables :

4) E1RATE = average rate of change of ENERG1, i.e. average rate of change in total energy consumption

5) E2RATE = average rate of change in ENERG2,i.e. average rate of change in electricity consumption

6) E3RATE = average rate of change in ENERG3,i.e. indicating trends in electricity-intensiveness relative to other fuel sources

Another resource utilisation variable worthy of testing would be some measure of labour skills employed by Tasmanian manufacturing industries. Ideally, it would have been desirable to have used some 'quality of labour' variable similar to Parry's SKILL1, i.e. the ratio of scientists, engineers, technicians, professional and administrative employees to production employees. Unfortunately, Tasmanian data was only available on working proprietors, on administrative and sales personnel plus distributors as a group, and on production employees including scientists, engineers, etc. Thus the 'quality of labour' variable that could be constructed was not expected to perform as well as could be desired, nevertheless, it provided some measure of the level of managerial skills present. Hence we have

7) LABQAL = the ratio of working proprietors and administrative personnel,etc. to production employees, etc. for industry (x)

Protection

Since the Tasmanian economy is dependent to some extent on world markets because of the limited size of its own domestic market, one could expect the more highly protected industries to be more insulated from fluctuations in external markets. Therefore a positive relationship could be expected between high rates of assistance and industries with 'good' performance. On the other hand, it is often argued that industries that are highly protected tend to become inefficient and stagnant because they are protected from competition from more efficient producers. Protection rates were therefore included in the analysis. Here again data limitations necessitated the use of available resources, so that the only protection variables that could be tested were :

8) EFRT78 = average effective rate of protection for Tasmanian industry (x) in 1977-78

9) NOUT78 = average nominal rate of protection for Tasmanian industry (x) in 1977-78

10) NMAT78 = average nominal rate of protection on materials for Tasmanian industry (x),1977-78

Since protection levels by broad industry categories tend in general to remain fairly constant over a period of time, it was felt that in the absence of other data, the application of these rates of assistance was justified.

Export Intensity

As we have seen in Chapter 2, it is generally believed that be it through the limitations of the home market or via the mechanisms of Wilde's core-periphery model, Tasmanian manufacturing is export-oriented. One should therefore consider the export-intensity of various industries an important structural characteristic. Although it would have been desirable to have used Trade and Shipping statistics to construct export-intensity variables time constraints did not permit the conversion of data under Trade and Shipping classifications to the corresponding ASIC classifications. It was therefore necessary to employ more readily available data from the 1977-78 input-output tables for Tasmania in order to derive the following variables :

11) OSEA78 = ratio of the value of exports overseas of industry (x) to the value of total supply

12) STAT78 = ratio of the value of exports interstate of industry (x) to the value of total supply

13) EXPO78 = ratio of the value of total exports of industry (x) to the value of total supply

This was not a very satisfactory set of variables because it could not account for changes in the level of exports over the time period under study. However, additional data were available in the form of 1968-69 input-output figures, so an average of these values were taken to result in a modified set of variables :

14) OSEAAV = average of OSEA78 and OSEA69

15) STATAV = average of STAT78 and STAT69

16) EXPOAV = average of EXPO78 and EXPO69

Although still far from perfect, this latter set of variables yielded noticeably improved results.

Finally, ABS data from trade and shipping statistics for Tasmania was used to construct a third set of export-intensity variables. However, inconsistencies in product classifications and confidentiality requirements resulted in the derived figures being virtually unusable.

Tasmanian-based versus Mainland-based

As concern is often expressed about local ownership the following variables attempt to reflect the extent or influence of Tasmanian ownership and/or control. The 'state of collection' refers to the state from which ABS data is collected, so presumably a firm's headquarters would be located in that state.

17) TASEST = ratio of the number of establishments whose state of collection is Tasmania to the total number of establishments operating in that industry in Tasmania

Thus TASEST reflects the proportion of establishments whose headquarters are located in Tasmania, and can therefore be considered Tasmanian-owned or -controlled.

18) TASEMP = ratio of employees of establishments whose state of collection is Tasmania to the total number of employees in that industry

Thus TASEMP attempts to reflect the employment aspect, or contribution to employment of Tasmanian-controlled industry.

CHAPTER 4. RESULTS

This chapter presents the results of four cluster analyses and the subsequent discriminant analyses applied to the clustered industries. The first section reports on the main cluster analysis undertaken on Tasmanian manufacturing industries over the time period 1975 to 1982. This is followed by two sections on the results of discriminant analysis and stepwise discriminant analysis on the various clustered groups.

The fourth section briefly discusses the consequences of clustering industries into only positive growth and negative growth clusters, and is followed by a section on discriminant analysis results on these two clusters.

Since the time period used for the first cluster analysis, 1975 to 1982, could be viewed as an atypical period due to the recessionary conditions prevailing not only in Australia but all over the world, cluster analysis was also applied to Tasmanian manufacturing industries for the time period 1968 to 1974 and the results are compared with those for 1975 to 1982 in section 6. Data and time constraints allowed only a brief examination of differences in the resulting groups via discriminant analysis and the results are presented in section 7.

Finally, a cluster analysis was attempted on the basis of two variables : growth in the period 1968 to 1974, and growth in the period 1975 to 1982. These results are discussed in the last section of this chapter.

4.1 CLUSTER ANALYSIS FOR 1975 TO 1982

Cluster analysis was undertaken on 32 industries, 5 fewer than the number of 3-digit ASIC manufacturing industries. The 5 excluded industries had GSP ratios in the neighbourhood of zero, i.e. their contribution to Tasmania's Gross State Product was negligible. These dropped industries were

214 Margarine, Oils & Fats
244 Knitting Mills
334 Photographic, Professional & Scientific Equipment
345 Leather & Leather Products and
346 Rubber Products

Thus 32 industries were left to be clustered. The industries were clustered on the basis of one performance variable, average growth rate in value-added, and the criterion selected for assigning individuals to clusters was the error sum of squares. This method is suitable for finding tight clusters which have the property that each cluster center represents the constituent cases at a high level of similarity with respect to all the underlying variables.

The 32 industries, or cases, were randomly assigned to 10 initial clusters with the most similar clusters fused together to form new cluster. This process was repeated until 3 terminal clusters were reached.

The results showed that the industries concerned could be clearly divided into 3 groups. The first group, or cluster, consisted of only 3 industries (323, 347 & 348). The minimum average growth rate of these 'high flyers' was 77.19 percent and the maximum was 85.72 percent. These three outstanding industries are

323 Motor Vehicles & Parts
347 Plastic & Related Products
348 Other Miscellaneous Manufacturing

The second cluster consisted of industries with only negative growth rates. There were 9 of these (212, 213, 218, 235, 294, 295-6, 316 & 335) and their average growth rates ranged from -28.28 percent to -3.03 percent. These 'declining' industries consisted of

212 Milk Products
213 Fruit & Vegetable Products
218 Beverages & Malt
235 Other Textile Products
253 Wood & Wood Products
294 Basic Iron & Steel
295-6 Non-Ferrous Metals & Non-Ferrous Metal Basic Products
316 Other Fabricated Metal Products
335 Appliances & Electrical Equipment

The remaining industries formed the last cluster, and their average growth rates ranged from 2.73 percent to 27.07 percent.

This particular partitioning of industries was unsatisfactory for two reasons. First, from an analytical point of view, the 'negative' cluster included both highly negative growth industries and those that were only slightly negative and may be regarded as stable industries. Similarly, the 'moderate growth' cluster included industries with relatively low and stable average growth rates. It would be undesirable to cluster the industries such that relatively stable industries are classified as either 'declining' or 'moderate growth' industries. Secondly, from a statistical viewpoint, there were large variances in the growth rates of the second and third clusters, indicating that the clusters were not as tight as they could perhaps be. Thus it was decided to go back one stage in the cluster analysis and examine the 4-cluster results.

At this stage, the industries fell into 4 satisfactory groupings: the first cluster was identical to that of the 3-cluster stage and comprised the 3 'high flyers'. The second cluster now consisted of 9 'moderate growth' industries, whose growth rates ranged from 11.93 percent to 27.07 percent. These industries are

211 Meat Products
215 Flour Mills & Cereal Food Products
217 Sugar & Other Food Products
234 Textile Fibres, Yarns & Woven Fabrics
276 Other Chemical Products
286 Clay Products & Refractories
287 Concrete & Concrete Products
288 Other Non-Metallic Mineral Products
324 Other Transport Equipment

The third cluster consisted of 13 industries, whose growth rates ranged from -3.98 percent to 6.38 percent. These may be considered 'stable' industries, and they are

212 Milk Products
216 Bread, Cakes & Biscuits
245 Clothing
253 Wood & Wood Products

254 Furniture & Mattresses
263 Paper & Paper Products
264 Printing & Allied Industries
275 Basic Chemicals
285 Glass & Glass Products
314 Fabricated Metal Products
315 Sheet Metal Products
336 Industrial Machinery & Equipment

The remaining 7 industries comprise the fourth cluster of clearly declining industries, whose growth rates ranged from -28.28 percent to -6.85 percent. These industries include

213 Fruit & Vegetable Products
218 Beverages & Malt
235 Other Textile Products
294 Basic Iron & Steel
295-6 Non-Ferrous Metals & Non-Ferrous Metal Basic Products
316 Other Fabricated Metal Products
335 Appliances & Electrical Equipment

At this point one may question why it is necessary to use a sophisticated clustering technique rather than a simple 'eyeballing' method. The reason for this is two-fold : first, a sophisticated clustering technique enables us to categorize individuals on a systematic basis, using a specified and precise mathematical criterion. Secondly, where more than one variable is used as the basis for clustering it can become a far less tractable task to handle manually.

There are two main diagnostic statistics that enable us to assess the validity of the resulting cluster, the F-ratio and the T-value. If we let

X_j = overall mean for variable (j)

S_j = overall standard deviation for variable (j)

V_j = overall variance of variable (j), i.e. $V_j = S_j^2$ and furthermore let the equivalent statistics for the subset of cases comprising a cluster (c) be denoted by X_{cj}, S_{cj} and V_{cj}, the F-ratio and the T-value are defined by Wishart (1978, p.77) as

F-ratio = V_{cj}/V_j

and T-value = $(X_{cj} - X_j)/S_j$

Small F-ratios indicate variables that have comparatively low variations within the cluster and are therefore good diagnostics. The expected value of the F-ratio is unity. On the other hand, large absolute values of T indicate continuous variables which have cluster means which are substantially different from the population sample means, and its expected value is zero. The results for the 4 clusters are given below :-

For E(F) = 1.0 and E(T) = 0.0,

CLUSTER	F-RATIO	T-VALUE
1 High Flyers	0.0305	2.6701
2 Moderate growth	0.0518	0.2996
3 Stable	0.0147	-0.3036
4 Declining	0.0707	-0.9658

Thus the F-ratios indicate that the chosen growth variable was an appropriate variable to use in the clustering of the 32 industries, while the T-values show that only the 'high flyers' had a group mean that was significantly different from that of the sample population.

4.2 DISCRIMINANT ANALYSIS : RESULTS FOR INDIVIDUAL VARIABLES

Having completed the task of clustering the 32 industries into 4 performance groups, discriminant analysis can be applied to determine what characteristics, if any, are statistically associated with each group. The objective of discriminant analysis here is to identify any characteristics or attributes which effectively distinguish between performance groups.

To demonstrate the nature of discriminant analysis each variable was first tested separately for evidence of discriminating power. This is analogous to using characteristics such as hair colour or skin colour separately to determine if an individual can be classified as (say) Mongoloid or Caucasian. Often it is not any one particular characteristic that emerges as a significant discriminator, but rather some particular combination(s) of characteristics, for example, both hair colour and skin colour, as well as (say) height. Poor individual performance of a variable as a discriminator therefore does not necessarily mean that they are not suitable for further discriminant analysis.

The results for each variable are summarised in Table 4.2. It is clear that the only variable to emerge as being a significant discriminator between the four groups is E1RATE, i.e. the rate of change in total energy usage. It was the only variable for which significant

TABLE 4.2. DISCRIMINANT ANALYSIS RESULTS FOR INDIVIDUAL VARIABLES

VARIABLE	WILKS'S LAMBDA (U)	CHI-SQUARE (X^2)	LEVEL OF SIGNIFICANCE	(a)
ENERG1	.9212	1.519	.678	34.38
ENERG2	.8806	2.353	.503	37.50
ENERG3	.8388	3.252	.354	31.25
LABQAL	.9223	0.485	.785	34.38
EFRT78	.8785	3.692	.297	37.50
NOUT78	.8688	3.727	.293	40.63
NMAT78	.9153	2.522	.471	12.50
OSEA78	.9674	0.531	.767	34.38
STAT78	.8777	2.987	.352	46.88
EXPO78	.9130	1.456	.483	46.88
TASEST	.9583	1.215	.749	21.88
TASEMP	.8702	3.964	.265	34.38
E1RATE	.7634	7.693	.053	46.88
E2RATE	.8372	5.065	.167	40.63
E3RATE	.9356	1.896	.594	28.13
OSEAAV	.9381	1.821	.610	18.75
STATAV	.9343	1.936	.586	25.00
EXPOAV	.9407	1.741	.628	28.13

(a) Percentage of cases correctly classified differences existed in the group means. The standardised coefficient of the discriminant function in all cases is unity because the absolute value of the coefficient shows the variable's relative contribution to the function, so that where there is only a single variable being tested, that variable contributes 100 percent to that function. An examination of the group means given below show that a relatively high value of E1RATE, i.e. that an industry which has been increasing its share in total usage of energy in the State, will tend to be classified as a 'high flyer' or 'moderate growth' industry; it is doing relatively well in terms of expanding its value-added. On the other hand, industries with a relatively low value of E1RATE will tend to be classified as 'stable' or 'declining'.

The group mean for E1RATE for each of the clusters is given below :

CLUSTER	E1RATE
1. High Flyers	7.27
2. Moderate Growth	3.38
3. Stable	-5.56
4. Declining	-0.57

The results therefore indicate that during the time period in question, industries which have relatively high growth rates in terms of value-added tend to be characterised by a relatively high growth rate in terms of energy usage.

The results of E1RATE also hint at why the average proportion variables ENERG1, ENERG2 and ENERG3 did not perform well as discriminators. It could well be that for energy usage, it is not the absolute amount of energy utilised by an industry that will distinguish it from another industry in terms of performance, but rather the rate at which it increases or decreases its energy consumption.

Another point to consider is that as output increases one would expect inputs to also increase. E1RATE and E2RATE, each representing the total change in one input, i.e. all fuels and electricity respectively, could therefore be expected to be positively linked with expanding industries and negatively linked with declining industries. E3RATE, however, represents relative changes in factor inputs and thus has more economic meaning because it is not correlated with changes in total output.

It is may be that the other variables may have performed better if a full set of data were available over the whole time period. Other discriminant analyses, discussed below, seem to indicate that when more information is incorporated into the variable they may emerge as having some discriminating power. Therefore it cannot be said that export intensity or protection levels, for example, have nothing to do with the performance level of a given industry. What can be said is that given the present definitions of the variables representing these characteristics, the groups means of these variables are not significantly different in a statistical sense, so that the variables do not discriminate between the four performance groups.

4.3 STEPWISE DISCRIMINANT ANALYSIS

4.3.1 STEPWISE DISCRIMINANT ANALYSIS : BETWEEN 4 PERFORMANCE GROUPS

Stepwise discriminant analysis was undertaken on all four performance groups using three sets of variables. The first set consisted of the 3 energy consumption levels ENERG1, ENERG2 and ENERG3, the quality of labour variable LABQAL, the 3 original protection variables EFRT78, NOUT78 and NMAT78, the 3 original export intensity variables OSEA78, STAT78 and EXPO78 and finally the 'Tasmanian elements' TASEST and TASEMP. The second set of variables comprised only the modified energy consumption variables E1RATE, E2RATE and E3RATE and the export intensity variables OSEAAV, STATAV and EXPOAV. Finally the modified variables were taken to replace their original versions in the first set, resulting in the third set of variables.

Table 4.3.1A summarises the results of stepwise discriminant analysis on the 4 performance groups with the three sets of variables. The variables which satisfied the minimum conditions and were selected for inclusion by the stepwise procedure are identified in

TABLE 4.3.1A DISCRIMINANT ANALYSIS ON ALL PERFORMANCE GROUPS

VARIABLES IN SET 1 \ FUNCTION NUMBER	STANDARDISED COEFFICIENTS (NUMBER OF FUNCTIONS = 2) 1	2
ENERG1	0.92	0.71
ENERG2	1.09	-0.42
ENERG3	-	-
LABQAL	-	-
EFRT78	-	-
NOUT78	-	-
NMAT78	-	-
OSEA78	-	-
STAT78	-	-
EXPO78	-	-
TASEST	-	-
TASEMP	-	-
CANONICAL CORRELATION COEFFICIENT	.69	.09
WILKS'S LAMBDA	.52	.99
CHI-SQUARE	6.25	.08
DEG. OF FREEDOM	4	1
LEVEL OF SIGNIFICANCE	.18	.77

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 46.88%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

TABLE 4.3.1A DISCRIMINANT ANALYSIS ON ALL PERFORMANCE GROUPS (CONT.)

VARIABLES IN SET 2 \ FUNCTION NUMBER	STANDARDISED COEFFICIENTS (NUMBER OF FUNCTIONS = 3) 1	2	3
E1RATE	2.63	0.02	1.10
E2RATE	-1.45	0.22	-1.67
E3RATE	-1.42	0.93	0.50
OSEAAV	-	-	-
STATAV	-0.07	0.87	0.14
EXPOAV	-	-	
CANONICAL CORRELATION COEFFICIENT	.65	.41	.31
WILKS'S LAMBDA	.43	.75	.90
CHI-SQUARE	22.84	7.74	2.69
DEG. OF FREEDOM	12	6	2
LEVEL OF SIGNIFICANCE	.03	.26	.26

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 62.50%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

TABLE 4.3.1A DISCRIMINANT ANALYSIS ON ALL PERFORMANCE GROUPS (CONT.)

VARIABLES IN SET 3 \ FUNCTION NUMBER	STANDARDISED COEFFICIENTS (NUMBER OF FUNCTIONS = 3) 1	2	3
E1RATE	-0.31	4.12	0.22
E2RATE	-0.76	-3.58	-0.59
E3RATE	-0.69	2.07	-0.91
LABQAL	-	-	-
EFRT78	0.45	-0.37	-0.64
NOUT78	-	-	-
NMAT78	-	-	-
OSEAAV	-	-	-
STATAV	0.07	-0.09	0.92
EXPOAV	-	-	-
TASEST	0.82	0.82	0.79
TASEMP	0.76	0.64	-0.43
CANONICAL CORRELATION COEFFICIENT	.78	.75	.67
WILKS'S LAMBDA	.09	.24	.54
CHI-SQUARE	34.40	20.97	8.78
DEG. OF FREEDOM	21	12	5
LEVEL OF SIGNIFICANCE	.03	.05	.12

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 71.88%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

TABLE 4.3.1B GROUP MEANS OF DISCRIMINATING VARIABLES ($\overline{Z}_{ik}$)

CLUSTER / VARIABLE	HIGH FLYERS	MODERATE GROWTH	STABLE	DECLINING
ENERG1	0.09	1.32	4.97	2.24
ENERG2	0.09	0.83	4.07	1.50
ENERG3	60.56	44.35	66.63	37.40
LABQAL	50.54	36.19	34.31	33.49
EFRT78	32.50	4.80	47.83	33.50
NOUT78	21.50	4.60	22.67	15.50
NMAT78	12.50	4.80	8.17	8.00
OSEA78	n.a.	9.08	5.73	1.12
STAT78	n.a.	13.62	30.27	40.82
EXPO78	n.a.	22.70	35.37	41.93
TASEST	72.38	59.72	68.31	74.72
TASEMP	38.26	13.94	22.45	7.35
E1RATE	7.27	3.38	-5.56	-0.57
E2RATE	2.72	4.66	-1.93	-0.73
E3RATE	-6.08	-1.31	-2.20	-1.06
OSEAAV	0.00	3.31	2.61	0.55
STATAV	0.00	14.22	15.39	19.42
EXPOAV	0.00	17.54	18.01	19.79
OSEATS	0.59	10.03	3.10	0.20
STATTS	0.99	11.36	5.37	0.16
EXPOTS	1.58	21.39	8.47	0.36

Table 4.3.1A. The absolute value of the standardised coefficients show the relative importance of each discriminating variable with respect to the other variables included in the function(s). The signs of the standardised coefficients merely serve to show the contribution of the variables to the form of the function(s). Given the variables which emerge as good discriminators, the group means (i.e. the $\bar{Z}_{ik}$'s) of the variables as given in Table 4.3.1B will indicate the direction of discrimination.

Recall from Section 3.2 (p.26) that the maximum number of unique discriminant functions that can be obtained is equal to either the total number of groups less one, or the number of discriminating variables, whichever is smaller. Therefore the maximum number of unique functions that can be derived when there are four industry groups is three. The reason for less than this maximum number of functions being obtained is that the second and/or third functions describe overlapping spaces, i.e. the functions are redundant and contain no statistically significant information about the nature of group differences.

In the first discriminant analysis only ENERG1 and ENERG3 'entered into the equations' which although not significant even at the 10 percent level are still worthy of some discussion in order to clarify the interpretation of discriminant analysis results.

From the standardised coefficients it can be seen that ENERG3 was relatively more important in the first function and that the reverse was true for the second function. The implication here is that industries which use electricity intensively relative to other energy sources tend to fall into the high performance groups (including the 'stable' industries). The 'declining' industries tend to be characterised by a low level of electricity usage relative to other fuels. These results also suggest that the 'high flyers' and 'moderate growth' industries are characterised by high electricity usage relative to other fuels, but at the same time they tend to use less energy relative to the other performance groups. This is not a surprising result, if we believe that Tasmania is endowed with a relative abundance of energy as a factor of production, in so far as industries employing a relatively cheap factor of production, i.e. electricity in comparison to other fuels, could be expected to perform better, all other things being equal. Furthermore, the fact that high flyers and the moderate growth industries appear to consume less energy in total is consistent with the nature of the industries which constitute the two groups. For example, approximately one third of the industries in the moderate growth cluster involve the processing of primary products, an activity which does not require as much energy as heavier industries such as iron and steel, fabricated metal products, etc.

The derived set of classification equations correctly classified only 46.88 percent of the cases. Note that with 4 groups, the a priori probability of any industry being correctly classified is 25 percent. From this percentage, and together with the information given by the three indicators of discriminating power, i.e. the canonical correlation coefficient, Wilk's lambda and the chi-square statistic it can be concluded that given this set of variables, ENERG1 and ENERG2 were the 'best' discriminators between the four industry performance groups, but their discriminating power is still low.

In order to test whether or not rates of change would discriminate between the groups the next stepwise discriminant analysis included the modified versions of the energy utilisation variables and export intensity variables. It was thought that in the case of energy consumption, rates of change may be more important characteristics than absolute levels (as indicated by the high significance of the rate of change in total energy usage variable, E1RATE). As to the export intensity variables, it was thought that additional information (i.e. to average over 1968-69 and 1977-78) would result in a truer reflection of an industry's export intensity. This approach proved to be a step in the right direction.

Out of the six variables tested, four were selected for inclusion in the 3 discriminant functions derived as shown in Table 4.3.1A, while OSEAAV and EXPOAV were dropped by the stepwise procedure. E2RATE and E1RATE were the most important variables in the first and third functions while E3RATE and STATAV were the most important variables in the second function. The chi-square statistics showed that only the first function contained significant discriminating power. Nevertheless, the derived set of classification equations correctly classified 62.50 percent of the industries.

An examination of the group means shown in Table 4.3.1B reveals that high flyers and moderate growth industries are characterised by relatively high increases in total energy consumption, relatively high increases in electricity consumption, but the negative values for E3RATE, i.e. the rate of increase in electricity consumption relative to other fuels, seems to suggest that as total energy consumption increases, firms tend to use other energy sources rather than electricity. As the converse is true for stable and declining industries, this result may reflect Callaghan's claim that Tasmania's water resources for hydro-electric power have reached their limits, in the sense that as electricity becomes more expensive relative to other fuels, firms will move away from electricity consumption to consumption of other fuels. The result that high flyers and moderate growth industries are characterised by high increases in total energy usage while stable and declining industries are characterised by decreases in total energy usage also makes sense when the definition of performance in the context of this study is considered. Since performance is measured in terms of growth, it is not surprising to find that expanding industries will tend to require more energy over time, all other things being equal.

The group means on the interstate export intensity variable, STATAV, show that the high flyers, or rather 2 of the high flying industries for which export figures were available, namely plastic and related products and other manufacturing, export virtually nothing interstate while the declining industries export on average 20 percent of their total production. This again appears to be a reasonable result because the high flyers tend to be import-competing industries rather than export-oriented, a factor which could possibly contribute to an industry being classified as a 'high flyer' in the first place, since the protection system allows it insulation from fluctuations in external markets.

Having tested the modified variables on their own and finding that considerable discriminating power existed in these variables, a third discriminant analysis was undertaken with the original set of variables, but with the modified energy consumption and export intensity variables replacing the originals. From Table 4.3.1A, it can be seen that 3 discriminant functions were derived with all 3 energy consumption variables, EFRT78, STATAV, TASEST and TASEMP. The first two functions were highly significant at the 3 percent and 5 percent levels respectively. The most important variables were TASEST and E1RATE in the first two functions respectively, and both E3RATE and STATAV for third function.

Since the role of the energy consumption variables and interstate export intensity has already been discussed, we can focus our attention on the remaining variables. The interpretation of the role of the 'local element' variables TASEST and TASEMP is ambiguous, for the group means show that both high flyers and declining industries comprise a high proportion of Tasmanian-based firms. Similarly, the role of EFRT78 is also difficult to interpret, since the group means show that high flyers and declining industries have similar, middle-range values, while the stable and the moderate growth industries are the extremes. It is in cases like this when more than 2 groups are involved that the results of discriminant analysis becomes difficult to interpret.

Of course the chances of finding structural characteristics which do discriminate and lend themselves to clearer interpretation are increased by applying discriminant analysis to different pairs of performance groups in order to highlight differences between them. In so far as policy-makers are keen to distinguish between certain types of groups such as the extremes, this approach is worth pursuing.

4.3.2 STEPWISE DISCRIMINANT ANALYSIS : BETWEEN PAIRS OF GROUPS

As we have seen, when discriminant analysis is applied to more than 2 groups the results can become ambiguous and difficult to interpret. In such cases, separate discriminant analysis can be applied to subsets of groups in order to highlight any differences that may exist between them. This section reports the results of separate discriminant analyses undertaken on pairs of performance groups using the first 3 sets of variables tested in the previous section.

In some cases it was necessary to exclude certain variables for which there were no figures available for any industries comprising a performance group, but unless otherwise stated, the full sets of variables will have been tested in each of the following discriminant analyses. In all cases only one discriminant function is derived per analysis because there are only 2 groups involved.

4.3.2.1 HIGH FLYERS & MODERATE GROWTH INDUSTRIES

Stepwise discriminant analysis was applied to determine if there were any significant differences between the outstanding 3 industries comprising the 'high flyers' group and the 9 industries comprising the 'moderate growth' group. The results are summarised in Table 4.3.2.1. In the first set of variables, the ones relating to export intensity were excluded because data was unavailable for the high flyers.

From the first set of variables TASEST, TASEMP, NMAT78, EFRT78 and ENERG1 were found to be good discriminators, in descending order of importance. The chi-square statistic was highly significant, indicating that there are real differences in the group means of these variables. From the group means given in Table 4.3.1B that the high flyers have a relatively greater proportion of Tasmanian-based establishments and that these firms employ a higher proportion of people working in that industry than do the Tasmanian-based firms in the moderate growth group. This result could be taken as very rough support for Wilde's view of locally-established versus 'filtered-down' industry. The high flyers are also characterised by a much higher average level of protection and by lower total energy consumption. This function correctly classified 10 (or 83.33 percent) of the 12 industries.

With the second set of variables E3RATE emerged as the most important discriminator, followed by E2RATE and STATAV. Table 4.3.1B strongly indicates that the high flyers are characterised by a slower rate of increase in electricity consumption relative to other fuels than the moderate growth industries, and by virtually no exports

TABLE 4.3.2.1 DISCRIMINANT ANALYSIS ON HIGH FLYERS & MODERATE GROWTH INDUSTRIES

VARIABLES IN SET 1	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=1)
ENERG1	-4.62
ENERG2	-
ENERG3	-
LABQAL	-
EFRT78	18.14
NOUT78	-
NMAT78	19.48
TASEST	31.31
TASEMP	27.14
CANONICAL CORRELATION COEFFICIENT	.99
WILKS'S LAMBDA	.001
CHI-SQUARE	21.23
DEG. OF FREEDOM	5
LEVEL OF SIGNIFICANCE	.001

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 83.33%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

TABLE 4.3.2.1 DISCRIMINANT ANALYSIS ON HIGH FLYERS & MODERATE GROWTH INDUSTRIES (CONT.)

VARIABLES IN SET 2	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=1)
E1RATE	-
E2RATE	0.78
E3RATE	1.06
OSEAAV	-
STATAV	0.71
EXPOAV	-
CANONICAL CORRELATION COEFFICIENT	.69
WILKS'S LAMBDA	.52
CHI-SQUARE	5.48
DEG. OF FREEDOM	3
LEVEL OF SIGNIFICANCE	.14

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 83.33%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

TABLE 4.3.2.1 DISCRIMINANT ANALYSIS ON HIGH FLYERS & MODERATE GROWTH INDUSTRIES (CONT.)

VARIABLES IN SET 3	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=1)
E1RATE	-
E2RATE	-
E3RATE	-
LABQAL	-
EFRT78	5.84
NOUT78	-
NMAT78	7.20
OSEAAV	-
STATAV	-
EXPOAV	-
TASEST	10.28
TASEMP	7.76
CANONICAL CORRELATION COEFFICIENT	.99
WILKS'S LAMBDA	.001
CHI-SQUARE	18.56
DEG. OF FREEDOM	4
LEVEL OF SIGNIFICANCE	.001

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 83.33%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

interstate. However, these results must be considered in the light of the chi-square statistic, which indicates that the observed differences in the group means of these variables are not sufficiently significant that we can draw any definite conclusions. Nevertheless, the function correctly classified 83.33 percent of the 12 industries. Considering the values of the canononical correlation coefficient and Wilks's Lambda, this high level of accuracy is probably due to chance.

In view of the results of the first 2 discriminant analyses between these 2 groups, the results obtained with the third set of variables could have been expected. The discriminant function included TASEST, TASEMP, NMAT78 and EFRT78, and the relative importance of each variable was totally consistent with the first discriminant analysis. Not surprisingly, the chi-square statistic yielded a similar level of significance and the function correctly classified 83.33 percent of the industries.

4.3.2.2 HIGH FLYERS & STABLE INDUSTRIES

If significant differences can exist in certain attributes of industries in 'close' performance groups, other real differences in certain attributes would be expected to exist between reasonable 'distinct' groups such as stable industries and high flyers. In this case only 2 sets of variables were tested; the first set excluding the export intensity variables because figures were unavailable for the high flyers, and the second set containing the modified versions of the energy consumption and export intensity variables. It was felt that because of the nature of the stepwise procedure in selecting variables from a given list of variables it would be redundant to continue to test the set of modified variables by themselves. Therefore further stepwise discriminant analyses were undertaken using only two sets of variables, the full set of original variables and the set including the modified variables unless otherwise specified. The results are summarised in Table 4.3.2.2.

In the first discriminant analysis 5 of the 9 variables were found to be good discriminators. The most important of these were NMAT78 and NOUT78, followed by LABQAL, TASEMP and ENERG2. This is a somewhat surprising result, for examination of the group means in Table 4.3.1B on the nominal protection variables shows that although the high flyers appear to have a relatively higher average nominal rate of protection on materials, the group means of the average nominal rate of protection on output are very close. A possible explanation for this is that this could actually be a reflection of differences in the average effective protection rate. If the effective rate of protection is given by

$$e = (o - xm)/(1 - x)$$

TABLE 4.3.2.2 DISCRIMINANT ANALYSIS ON HIGH FLYERS & STABLE INDUSTRIES

VARIABLES IN SET 1	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=1)
ENERG1	-
ENERG2	0.59
ENERG3	-
LABQAL	2.57
EFRT78	-
NOUT78	-3.38
NMAT78	4.96
TASEST	-
TASEMP	-1.25
CANONICAL CORRELATION COEFFICIENT	.93
WILKS'S LAMBDA	.14
CHI-SQUARE	14.52
DEG. OF FREEDOM	5
LEVEL OF SIGNIFICANCE	.01

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 93.75%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

TABLE 4.3.2.2 DISCRIMINANT ANALYSIS ON HIGH FLYERS & STABLE INDUSTRIES (CONT.)

VARIABLES IN SET 2	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=1)
E1RATE	1.99
E2RATE	-
E3RATE	-
LABQAL	-
EFRT78	-
NOUT78	-
NMAT78	1.81
OSEAAV	-
STATAV	-
EXPOAV	-
TASEST	-0.52
TASEMP	-
CANONICAL CORRELATION COEFFICIENT	.92
WILKS'S LAMBDA	.15
CHI-SQUARE	16.09
DEG. OF FREEDOM	3
LEVEL OF SIGNIFICANCE	.001

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 87.50%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

where o = nominal rate on output

m = nominal rate on materials

and x = ratio of value of materials to value of output is assumed to be $0 < x < 1$

then, given the same nominal rate of protection on output and similar materials to output ratios, the effective rate will be lower the higher the nominal rate on materials. In other words, given a nominal rate on output, the effective rate will decrease as the nominal rate on materials increases. Thus the above result would imply that high flyers are characterised by a lower average effective rate of protection than stable industries. Indeed, the value for EFRT78 is lower for high flyers than the corresponding value for the stable group.

Table 4.3.1B also shows that in comparison to the stable industries, high flyers tend to have a higher ratio of administrative personnel to production workers and a lower level of electricity consumption. The results also show that of the Tasmanian-based firms in the 2 groups of industries, the firms in the 'high flyers' group employed a higher proportion of people than did firms in the 'moderate growth' group. The function was highly accurate in correctly classifying 93.75 percent of the 16 industries, but this high degree of accuracy is likely to be due to chance, since the chi-square statistic indicated non-significant differences in the group means of the variables.

In the second set of variables only 3 emerged as significant discriminators : ElRATE, NMAT78 and TASEST in descending order of importance. The group means of these variables in Table 4.3.1B show that the high flyers are characterised by a high rate of increase in total energy consumption while the stable industries are characterised by a high rate of decrease in total energy consumption. This result is consistent with the view that since performance is measured in terms of growth, the expanding industries would therefore be more likely to be those that have an increasing demand for energy. The results further indicate that the high flyers generally have a higher nominal rate of protection on material inputs than the stable industries and that they comprise a higher proportion of Tasmanian-based firms. This function was highly significant, and on the basis of the 3 selected variables, correctly classified 87.50 percent of the 16 industries.

4.3.2.3 HIGH FLYERS & DECLINING INDUSTRIES

This is perhaps the most interesting of all discriminant analyses to be undertaken between pairs of groups, for it attempts to highlight the differences between the two extremes in terms of performance : the high flyers and the declining industries. Not surprisingly, discriminant analysis with both sets of variables yielded highly significant results as shown in Table 4.3.2.3.

In the first set of variables, the export intensity variables OSEA78, STAT78 and EXPO78 were again excluded because no data was available for any of the high flyers. Out of the remaining 9 variables 3 were 'entered into the equation' so that the derived discriminant function included LABQAL, NOUT78 and TASEMP in descending order of importance. The group means on these variables as given in Table 4.3.1B show that in comparison to the declining industries the high flyers have a greater proportion of administrative personnel to production employees, perhaps indicating that high flying industries are relatively more capital-intensive than the declining industries. Given the present variables, no conclusions can be drawn on this point, but this is certainly an area for investigation in future research. The results further indicate that high flyers are afforded a higher nominal rate of protection on output than are the declining industries, and that the Tasmanian-based firms in the high flyers group in general account for a very much higher proportion of employment within each industry. The chi-square statistic showed significant differences in the group means of these 3 variables, resulting in a function which correctly classified 80.00 percent of the 10 industries.

As in the first discriminant analysis, the

TABLE 4.3.2.3 DISCRIMINANT ANALYSIS ON HIGH FLYERS & DECLINING INDUSTRIES

VARIABLES IN SET 1	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=1)
ENERG1	-
ENERG2	-
ENERG3	-
LABQAL	9.94
EFRT78	-
NOUT78	-7.24
NMAT78	-
TASEST	-
TASEMP	5.41
CANONICAL CORRELATION COEFFICIENT	.98
WILKS'S LAMBDA	.04
CHI-SQUARE	8.26
DEG. OF FREEDOM	3
LEVEL OF SIGNIFICANCE	.04

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 80.00%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

TABLE 4.3.2.3 DISCRIMINANT ANALYSIS ON HIGH FLYERS & DECLINING INDUSTRIES (CONT.)

VARIABLES IN SET 2	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=1)
E1RATE	-
E2RATE	-
E3RATE	14.18
LABQAL	-6.01
EFRT78	-
NOUT78	-
NMAT78	-
OSEAAV	9.61
STATAV	-
EXPOAV	-
TASEST	-
TASEMP	-16.25
CANONICAL CORRELATION COEFFICIENT	.99
WILKS'S LAMBDA	.001
CHI-SQUARE	13.23
DEG. OF FREEDOM	4
LEVEL OF SIGNIFICANCE	.01

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 90.00%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

second experiment including the modified variables found TASEMP and LABQAL to be good discriminators. The other 2 variables included in the discriminant function were E3RATE and OSEAAV. An examination of the group means of these 4 variables in Table 4.3.1B reveal that the Tasmanian-based establishments in the high flyers group of industries accounted for a much higher percentage of employment within the industries than did Tasmanian-based firms in the declining group. The group means also show that high flyers are characterised by a relatively higher rate of decrease in electricity consumption compared to other fuels and a higher ratio of administrative personnel to production employees. It should be noted that the role of the export intensity variable OSEAAV is difficult to interpret since the group means are very similar : 0.00 for the high flyers and 0.55 for the declining industries. Nevertheless, the observed differences on these variables between the 2 performance groups are shown to be highly significant at the one percent level. It is therefore unlikely to be by coincidence that the derived function was highly accurate in correctly classifying 90.00 percent of the industries.

4.3.2.4 MODERATE GROWTH INDUSTRIES & STABLE INDUSTRIES

Having determined certain significant differences between high flyers and all other groups of industries, discriminant analysis is then applied to identify any differences which may exist between the other performance groups. In this section moderate growth industries are compared to the stable industries, and for this purpose both sets of variables as listed in Section 4.3.2.3 are used. Table 4.3.2.4 presents the results for the discriminant analyses.

In the first discriminant analysis, the chi-square statistic indicated that there were no significant differences between the group means of the discriminating variables, but since 2 of the variables reappear in the second discriminant analysis, it is worth commenting briefly on some of these differences. The variables which qualified as discriminators were ENERG2, EFRT78, TASEST and LABQAL in descending order of importance. Comparison of the group means in Table 4.3.1B shows that the stable industries are characterised by a higher level of electricity consumption than the moderate growth industries, and that the average effective rate of protection is much higher. The stable industries also appear to have a greater proportion of Tasmanian-based firms. Once again the role of LABQAL is difficult to interpret because the group means are so close, with the value for moderate growth industries being slightly higher than that for the stable industries. The derived function correctly classified 72.73 percent of the 22 industries.

TABLE 4.3.2.4 DISCRIMINANT ANALYSIS ON MODERATE GROWTH & STABLE INDUSTRIES

VARIABLES IN SET 1	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=1)
ENERG1	-
ENERG2	1.37
ENERG3	-
LABQAL	0.79
EFRT78	1.13
NOUT78	-
NMAT78	-
OSEA78	-
STAT78	-
EXPO78	-
TASEST	0.95
TASEMP	-
CANONICAL CORRELATION COEFFICIENT	.79
WILKS'S LAMBDA	.38
CHI-SQUARE	6.73
DEG. OF FREEDOM	5
LEVEL OF SIGNIFICANCE	.15

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 72.73%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

TABLE 4.3.2.4 DISCRIMINANT ANALYSIS ON MODERATE GROWTH & STABLE INDUSTRIES (CONT.)

VARIABLES IN SET 2	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=1)
E1RATE	1.58
E2RATE	-
E3RATE	1.04
LABQAL	-0.83
EFRT78	-
NOUT78	-
NMAT78	-
OSEAAV	-
STATAV	-
EXPOAV	-
TASEST	-1.11
TASEMP	-
CANONICAL CORRELATION COEFFICIENT	.81
WILKS'S LAMBDA	.34
CHI-SQUARE	11.81
DEG. OF FREEDOM	4
LEVEL OF SIGNIFICANCE	.02

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 72.73%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

The second discriminant analysis identified 4 variables as discriminators : E1RATE, TASEST, E3RATE and LABQAL in descending order of importance. From the group means given in Table 4.3.1B it can be seen that moderate growth industries have a positive rate of increase in total energy consumption, comprise a lower proportion of Tasmanian-based firms than do stable industries, and have in general a slightly higher proportion of administrative personnel to production employees. The stable industries, on the other hand, are characterised by decreases in total energy consumption, and in both groups there appears to be a trend away from electricity consumption in favour of other fuel sources. On the basis of these variables, the function correctly classified only 72.73 percent of the cases. One possible explanation for this relatively low classification performance of the discriminant function is that the groups may be very close so that a more distinct separation can be obtained if a greater range of variables were available.

4.3.2.5 MODERATE GROWTH INDUSTRIES & DECLINING INDUSTRIES

Discriminant analysis was applied using 2 sets of variables, and highly significant differences were found in both experiments. The results are summarised in Table 4.3.2.5.

The first discriminant analysis identified EFRT78, NOUT78, EXPO78 and OSEA78 as good discriminators between moderate growth industries and declining industries. The group means for these variables, given in Table 4.3.1B, show that the declining industries have a far higher average effective rate of protection, mainly because of the higher nominal rate on their output. The declining industries also export a higher proportion of their output, but a lower percentage of these exports go to overseas destinations than those of the moderate growth industries. This result would seem to support the theory that because protection raises the domestic cost structure, export activities are disadvantaged so that manufacturers will tend to concentrate more on the protected home market, in this case, Australia as a whole. The products of moderate growth industries, with their lower average effective rate of protection, would thus be more competitive in world markets, resulting in the relatively higher proportion of exports to overseas destinations. The chi-square statistic indicates that these differences in the discriminating variables are highly significant. However, the function correctly classified only 68.78 percent of the 20 industries, suggesting that although the differences in the selected variables were significant there may have been other important variables which were not included in the original set, due , for example, to lack of available data.

The second discriminant analysis was slightly

TABLE 4.3.2.5 DISCRIMINANT ANALYSIS ON MODERATE GROWTH & DECLINING INDUSTRIES

VARIABLES IN SET 1	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=1)
ENERG1	-
ENERG2	-
ENERG3	-
LABQAL	-
EFRT78	11.78
NOUT78	-7.86
NMAT78	-
OSEA78	-2.61
STAT78	-
EXPO78	6.30
TASEST	-
TASEMP	-
CANONICAL CORRELATION COEFFICIENT	.99
WILKS'S LAMBDA	.02
CHI-SQUARE	11.95
DEG. OF FREEDOM	4
LEVEL OF SIGNIFICANCE	.02

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 68.78%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

TABLE 4.3.2.5 DISCRIMINANT ANALYSIS ON MODERATE GROWTH & DECLINING INDUSTRIES (CONT.)

VARIABLES IN SET 2	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=1)
E1RATE	-
E2RATE	-13.55
E3RATE	-
LABQAL	30.38
EFRT78	-
NOUT78	-24.92
NMAT78	-
OSEAAV	13.05
STATAV	4.39
EXPOAV	-
TASEST	-
TASEMP	-
CANONICAL CORRELATION COEFFICIENT	.99
WILKS'S LAMBDA	.002
CHI-SQUARE	28.20
DEG. OF FREEDOM	5
LEVEL OF SIGNIFICANCE	.001

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 75.00%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

more successful, in that given highly significant differences in the discriminating variables, the function correctly classified 75.00 percent of the industries. The variables which qualified as discriminators were LABQAL, NOUT78, E2RATE, OSEAAV and STATAV in descending order of importance. Table 4.3.1B shows that given these variables the moderate growth industries are shown to have a higher administrative to production personnel ratio, be less protected and export a higher percentage of total exports to overseas destinations. The declining industries tend to export a higher percentage of their total exports to other Australian States. With respect to the rate of electricity consumption, the moderate growth industries are characterised by a rate of increase of around 5 percent per annum, while the electricity consumption of the declining industries on the whole remains stable.

4.3.2.6 STABLE INDUSTRIES & DECLINING INDUSTRIES

The results of discriminant analysis applied to stable industries and declining industries are perhaps the most surprising and interesting of all comparisons so far. It was expected that discriminant analysis would identify significant differences in certain characteristics between distinct groups such as high flyers and declining industries, but the results obtained in this section show that it is also possible to find variables which appear to distinguish with a very high degree of accuracy between 'close' groups such as stable industries and declining industries. As before, the 2 sets of variables were tested, and the results are summarised in Table 4.3.2.6.

In the first discriminant analysis, ENERG3 was identified as the most important discriminator, followed by NMAT78, OSEA78, TASEST, TASEMP and ENERG2. Examination of the group means in Table 4.3.1B reveals that stable industries tend to consume more electricity than do the declining industries, and that they are more electricity-intensive in relation to other energy sources. This could be due to the nature of the industries comprising the declining group, where 3 out of the 7 include activities such as fruit and vegetable products which would not be heavy users of electricity. Like the moderate growth industries, the stable industries are also characterised by a relatively higher level of exports overseas, reflecting a greater degree of competitiveness in world markets than the declining industries. With respect to the 'local element', the stable industries comprise a lower percentage of Tasmanian-based firms, yet these firms account for a much higher proportion of employment within those industries. The derived function correctly classified only 60 percent of the industries, indicating that although the differences in the discriminating variables were significant, they were not sufficient to clearly distinguish between the 2 performance groups.

TABLE 4.3.2.6 DISCRIMINANT ANALYSIS ON STABLE INDUSTRIES & DECLINING INDUSTRIES

VARIABLES IN SET 1	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=1)
ENERG1	-
ENERG2	1.81
ENERG3	16.40
LABQAL	-
EFRT78	-
NOUT78	-
NMAT78	-7.21
OSEA78	7.11
STAT78	-
EXPO78	-
TASEST	-6.07
TASEMP	-3.53
CANONICAL CORRELATION COEFFICIENT	.99
WILKS'S LAMBDA	.01
CHI-SQUARE	14.87
DEG. OF FREEDOM	6
LEVEL OF SIGNIFICANCE	.02

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 60.00%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

TABLE 4.3.2.6 DISCRIMINANT ANALYSIS ON STABLE INDUSTRIES & DECLINING INDUSTRIES (CONT.)

VARIABLES IN SET 2	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=1)
E1RATE	5.63
E2RATE	-7.24
E3RATE	7.36
LABQAL	2.32
EFRT78	-3.39
NOUT78	-
NMAT78	-
OSEAAV	-0.75
STATAV	-
EXPOAV	-
TASEST	-
TASEMP	-
CANONICAL CORRELATION COEFFICIENT	.94
WILKS'S LAMBDA	.12
CHI-SQUARE	19.21
DEG. OF FREEDOM	6
LEVEL OF SIGNIFICANCE	.004

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 95.00%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

In the second discriminant analysis all 3 energy consumption variables E1RATE, E2RATE & E3RATE qualified as discriminators together with LABQAL, EFRT78 and OSEAAV. The group means given in Table 4.3.1B show that while the energy and electricity consumption of the declining industries remained at relatively stable levels, the stable industries were characterised by decreasing total energy consumption with the decrease mainly in electricity. Table 4.3.1B also shows that the interpretation of LABQAL as a discriminator is difficult because of the very similar group means. Nevertheless it may be noted that the ratio of administrative personnel to production workers is only slightly higher for stable industries than for declining industries. One seemingly inconsistent result is that although the stable industries are on average more highly protected than the declining industries, they appear to export a somewhat higher proportion of their total exports to overseas destinations. This result, however, could be due to aggregation and to the problems associated with the estimation of the average export intensity variables. Nevertheless, on the basis of these variables, 94 percent of the industries were correctly classified, and the chi-square statistic indicates that this high degree of accuracy is the result of highly significant differences between the group means of the discriminating variables selected for inclusion in the discriminant function.

4.4. CLUSTERING INDUSTRIES INTO POSITIVE & NEGATIVE GROWTH CLUSTERS

As seen in the previous sections, the results of discriminant analysis are dependent not only upon the variables chosen for testing but also on the clustering of the industries themselves. Therefore further different clusterings of industries were attempted to see if other significant discriminating variables could be identified. The first cluster variation to be attempted was to simply divide the industries into two groups : 1) those which had a positive average growth rate from 1975 to 1982 and 2) those which had a negative average growth rate for that period.

The 'positive growth' cluster included 23 manufacturing industries :

211 Meat Products
215 Flour Mills & Cereal Food Products
216 Bread, Cakes & Biscuits
217 Sugar & Other Food Products
234 Textile Fibres, Yarns & Woven Fabrics
245 Clothing
246 Footwear
254 Furniture & Mattresses
263 Paper & Paper Products
264 Printing & Allied Industries
275 Basic Chemicals
276 Other Chemical Products
285 Glass & Glass Products
286 Clay Products & Refractories
287 Concrete & Concrete Products
288 Other Non-Metallic Mineral Products
314 Fabricated Metal Products
315 Sheet Metal Products
323 Motor Vehicles & Parts
324 Other Transport Equipment
336 Industrial Machinery & Equipment
347 Plastics & Related Products
348 Other Miscellaneous Manufacturing

The average growth rates of these industries ranged from 2.73 percent for industrial machinery and equipment to 85.72 percent for plastic and related products.

The remaining 9 industries listed below comprised the second cluster, i.e. the 'negative growth' industries :

212 Milk Products
213 Fruit & Vegetable Products
218 Beverages & Malt
235 Other Textile Products
253 Wood & Wood Products
294 Basic Iron & Steel
295-6 Non-Ferrous Metals & Non-Ferrous Metal Basic Products
316 Other Fabricated Metal Products
335 Appliances & Electrical Equipment

The average growth rates of these industries vary from -3.03 percent for wood & wood products to -28.28 percent for other textile products.

The intention here is not to divide industries into 'good' performers and 'bad' performers, but rather to provide a basis for identifying any differences which may exist in characteristics between industries with a positive average growth rate and those with a negative average growth rate.

4.5. DISCRIMINANT ANALYSIS FOR POSITIVE AND NEGATIVE GROWTH CLUSTERS

The 32 manufacturing industries used in the first cluster analysis for the period 1975 to 1982 were divided into to 2 groups : those with positive average growth rates and those with negative average growth rates. Discriminant analysis was then applied to determine if any significant differences exist in certain attributes of these industries. As in the case of the first clustering of industries, the variables were first tested one by one for evidence of individual discriminating power. Next a stepwise discriminant analysis was undertaken in search of some combination(s) of variables which may distinguish between the two groups.

4.5.1 DISCRIMINANT ANALYSIS : RESULTS FOR INDIVIDUAL VARIABLES

Not all variables were individually tested because this application of discriminant analysis only serves to provide a rough indication of what discriminating power may exist. All variables, however, were used in the more important stepwise discriminant analyses.

The variables which were tested on the basis of this simple clustering were : ENERG1, ENERG2, ENERG3, EFRT78, NOUT78, OSEA78, STAT78, EXPO78, TASEST and TASEMP. The results are summarised in Table 4.5.1A. It is obvious from the table that none of the variables performed well as discriminators between the two groups. The only variable that may be of interest is the

TABLE 4.5.1A DISCRIMINANT ANALYSIS RESULTS FOR INDIVIDUAL VARIABLES

VARIABLE	WILKS'S LAMBDA (U)	CHI-SQUARE (X^2)	LEVEL OF SIGNIFICANCE	(a)
ENERG1	.9995	0.009	.924	43.75
ENERG2	.9999	0.001	.970	46.88
ENERG3	.9681	0.633	.426	56.25
EFRT78	.9696	0.910	.340	50.00
NOUT78	.9542	1.290	.256	53.13
OSEA78	.9867	0.221	.638	65.63
STAT78	.9570	0.725	.394	59.38
EXPO78	.9356	1.098	.295	65.63
TASEST	.9991	0.026	.873	43.75
TASEMP	.9142	2.647	.104	46.88

(a) percentage of cases correctly classified

TABLE 4.5.1B GROUP MEANS OF DISCRIMINATING VARIABLES ($\bar{Z}_{ik}$)

CLUSTER / VARIABLE	POSITIVE GROWTH	NEGATIVE GROWTH
ENERG1	3.40	2.29
ENERG2	2.71	1.47
ENERG3	58.60	36.80
LABQAL	34.98	34.65
EFRT78	29.10	29.00
NOUT78	15.50	11.67
NMAT78	7.20	5.67
OSEA78	5.69	8.37
STAT78	23.70	31.45
EXPO78	29.01	39.81
TASEST	62.95	76.13
TASEMP	20.44	4.90
E1RATE	-1.21	-2.22
E2RATE	1.51	-1.28
E3RATE	-2.73	-2.57
OSEAAV	3.38	2.48
STATAV	16.90	19.95
EXPOAV	20.22	22.22

Tasmanian employment variable, TASEMP, for which was found a significant difference in the group means at the 10 percent level. The group means for all variables is given in Table 4.5.1B.

4.5.2 STEPWISE DISCRIMINANT ANALYSIS

Having obtained no significant evidence of discriminating power in individual variables, the next step was to apply stepwise discriminant analysis to the positive growth industries and the negative growth industries using full sets of variables. The first set of variables included ENERG1, ENERG2, ENERG3, LABQAL, EFRT78, NOUT78, NMAT78, OSEA78, STAT78, EXPO78, TASEST and TASEMP. The second set of variables was similar to the first, but the energy consumption variables were replaced by E1RATE, E2RATE and E3RATE, and the export intensity variables were replaced by OSEAAV, STATAV and EXPOAV. The results are summarised in Table 4.5.2.

In the first discriminant analysis, 5 variables qualified as discriminators : STAT78, TASEST, ENERG3, ENERG2 and LABQAL. However, the chi-square statistic indicated that the differences in the group means of these variables as given in Table 4.5.1B were not real differences and consequently the function correctly classified only 65.63 percent of the industries.

The second set of variables also failed to produce a set of good discriminators, for again the chi-

TABLE 4.5.2 DISCRIMINANT ANALYSIS ON POSITIVE GROWTH & NEGATIVE GROWTH CLUSTERS

VARIABLES IN SET 1	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=1)
ENERG1	-
ENERG2	-0.98
ENERG3	-1.06
LABQAL	0.68
EFRT78	-
NOUT78	-
NMAT78	-
OSEA78	-
STAT78	2.17
EXPO78	-
TASEST	2.13
TASEMP	-
CANONICAL CORRELATION COEFFICIENT	.80
WILKS'S LAMBDA	.36
CHI-SQUARE	8.70
DEG. OF FREEDOM	5
LEVEL OF SIGNIFICANCE	.12

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 65.63%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

TABLE 4.5.2 DISCRIMINANT ANALYSIS ON POSITIVE GROWTH & NEGATIVE GROWTH CLUSTERS (CONT.)

VARIABLES IN SET 2	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=1)
E1RATE	-
E2RATE	0.68
E3RATE	-
LABQAL	-
EFRT78	0.83
NOUT78	-
NMAT78	-
OSEAAV	-
STATAV	-
EXPOAV	-
TASEST	-0.82
TASEMP	0.57
CANONICAL CORRELATION COEFFICIENT	.57
WILKS'S LAMBDA	.68
CHI-SQUARE	6.65
DEG. OF FREEDOM	4
LEVEL OF SIGNIFICANCE	.16

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 65.63%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

square statistic indicated no significant differences in the group means of the qualifying variables. The function, which contained EFRT78, TASEST, E2RATE and TASEMP, also correctly classified only 65.63 percent of the industries.

It is not surprising that discriminant analysis between positive growth and negative growth industries did not yield any significant results. Although manufacturing industries can be clearly divided into positive growth and negative growth clusters, such a division is not practical from an analytical viewpoint since the difference between the slowest expanding industry and the slowest contracting industry in terms of value-added is very small. Furthermore, industries within such broad clusters cannot be expected to have a high degree of homogeneity, and as such, can be expected to vary greatly in their characteristics. Nevertheless, the possibility that some discriminating attributes do exist should not be discredited until further investigation into this area is carried out.

4.6. CLUSTER ANALYSIS FOR 1968 TO 1974

Cluster analysis has been undertaken on 3-digit ASIC industries for Tasmanian manufacturing over the period 1975 to 1982. Since this period may be considered atypical in that it is a recessionary period, and thus the associations one has been able to find are only relevant to such a period, further cluster analysis was undertaken on Tasmanian manufacturing industries at the 3-digit ASIC level for an earlier period, i.e. 1968 to 1974. The procedure for the estimation of value-added was consistent with that previously used, so that in the end 32 industries remained to be clustered. As before, the industries were clustered on the basis of one performance variable, average growth or the rate of change in value-added, and the criterion selected for assigning individuals to clusters was the error sum of squares.

The 32 industries were randomly assigned to 10 initial clusters and were then relocated one at a time to different clusters with the most similar clusters fused together to form new clusters. This process was repeated until 3 terminal clusters were reached.

The best results were obtained with 5 clusters of industries. The first cluster consisted of only one industry, plastic and related products, with an outstanding average growth rate of 152.34 percent. The second cluster comprised 5 other 'high flyers' :

217 Sugar & Other Food Products
245 Clothing
294 Basic Iron & Steel
324 Other Transport Equipment
348 Other Miscellaneous Manufacturing

with average growth rates ranging from 34.87 percent to 57.05 percent.

The third cluster consisted of 6 'moderate growth' industries, whose average growth rates ranged from 8.40 percent to 21.20 percent. These industries were

211 Meat Products
215 Flour Mills & Cereal Food Products
235 Other Textile Products
288 Other Non-Metallic Mineral Products
336 Industrial Machinery & Equipment

The fourth and largest cluster included 11 industries which could be considered 'stable' and these were

212 Milk Products
218 Beverages & Malt
234 Textile Fibres, Yarns & Woven Fabrics
246 Footwear
254 Furniture & Mattresses
263 Paper & Paper Products
264 Printing & Allied Industries
275 Basic Chemicals
285 Glass & Glass Products
287 Concrete & Concrete Products
314 Fabricated Metal Products

The growth rates of these industries ranged between -4.28 percent and 3.26 percent.

Finally, the remaining 9 industries formed the fifth cluster, which may be called the 'declining' industries. The growth rates of these industries ranged from -16.92 percent to -5.71, and this group consisted of

213 Fruit & Vegetable Products
216 Bread, Cakes & Biscuits
276 Other Chemical Products
286 Clay products & Refractories
295-6 Non-Ferrous Metals & Non-Ferrous Metal Basic Products
315 Sheet Metal Products
316 Other Fabricated Metal Products
323 Motor Vehicles & Parts
335 Appliances & Electrical Equipment

To assess the validity of these results the F-ratios and the T-values for each cluster is examined. (See page 56 above). Small F-ratios indicate variables that have comparatively low variations within the cluster and are therefore good diagnostics. The expected value of the F-ratio is unity. Large T-values indicate continuous variables which have cluster means that are substantially different from the population sample means, and the expected valued is zero. The results for the 5 clusters are given below :

For E(F)=1.0 and E(T)=0.0,

CLUSTER	F-RATIO	T-VALUE
1 Outstanding	0.0000	4.4417
2 High Flyers	0.0809	1.0234
3 Moderate Growth	0.0220	0.0608
4 Stable	0.0065	-0.3596
5 Declining	0.0167	-0.6631

The F-ratios for all clusters indicated that the chosen growth variable was an appropriate one to use in the clustering of the 32 industries, while the T-values show that only the outstanding industry and the high flyers had group means which were significantly different from that of the sample population.

For practical purposes it seemed appropriate to include plastic and related products with the 5 high flyers to form one 'high flyers' cluster so that a comparison could be made between the 4 performance clusters in each time period. Table 4.6 lists the 32 industries and summarises the comparisons between the 2 time periods.

It can be seen from Table 4.6 that out of the 32 industries 17 remained in the same clusters over both periods. Of the 15 that had changed clusters over the 2

TABLE 4.6 COMPARISON OF CLUSTERED INDUSTRIES FOR THE PERIODS 1968-1974 & 1975-1982

PERIOD / ASIC	1968-74		1975-82
211	2		2
212	3		3
213	4		4
215	2		2
216	4	>>>	3
217	1	<<<	2
218	3	<<<	4
234	3	>>>	2
235	2	<<<	4
245	1	<<<	3
246	3		3
253	2	<<<	3
254	3		3
263	3		3
264	3		3
275	3		3

1 = high flyer
2 = moderate growth industry
3 = stable industry
4 = declining industry
<<< to 'lower' group
>>> to 'higher' group

TABLE 4.6 COMPARISON OF CLUSTERED INDUSTRIES FOR THE PERIODS 1968-74 & 1975-82 (CONT.)

PERIOD / ASIC	1968-74		1975-82
276	4	>>>	2
285	3		3
286	4	>>>	2
287	3	>>>	2
288	2		2
294	1	<<<	4
295-6	4		4
314	3		3
315	4	>>>	3
316	4		4
323	4	>>>	1
324	1	<<<	2
335	4		4
336	2	<<<	3
347	1		1
348	1		1

1 = high flyer
2 = moderate growth industry
3 = stable industry
4 = declining industry
<<< to 'lower' group
>>> to 'higher' group

periods, 7 had moved to higher performance clusters and 8 had moved to lower performance clusters. Although most industries moved only one step either way in terms of performance groups, several saw substantial changes in their performance status.

The most noticeable change occurred in the basic iron and steel industry, which went from being in the high flyers group of the earlier period to the declining group of the recessionary period. The clothing industry was another high flyer which was classified as 'stable' in the second period, while other textile products changed from a moderate growth to a declining industry over the two periods. On the other hand, changes in industry performance had also gone in the opposite direction. The most remarkable of these is the motor vehicles and parts industry, which went from being a declining industry to a high flyer. There were 2 other industries which changed from the declining group to the moderate growth group. These industries were other chemical products and clay products and refractories. Thus given these changes in the composition of the clusters, discriminant analysis could be expected to yield different results when applied to these new clusters.

4.7. DISCRIMINANT ANALYSIS FOR 1968 TO 1974 CLUSTERS

Now that the 32 industries have been clustered into 4 performance groups for the years 1968 to 1974, it would be desirable to apply discriminant analysis in the same manner as previously done for the 1975 to 1982 clusters so that the results can be compared. Unfortunately, extremely limited data was available, in addition to which time constraints permitted only the currently available data to be used for testing. Thus the first set of variables consisted of only ENERG1, ENERG2, ENERG3 and LABQAL while the second set consisted of E1RATE, E2RATE, E3RATE and LABQAL. Table 4.7A summarises the results of the stepwise discriminant analysis performed on these 2 sets of variables.

In the first discriminant analysis, no variables qualified for the analysis. This is not as surprising a result as it may at first seem, because this could merely be due to the possibility that none of the energy consumption variables have sufficient discriminating power without the presence of other strong discriminators.

In the second discriminant analysis 3 functions were derived which included all 3 rate of energy consumption variables. However, the values for chi-square indicate that only the first function was significant. On the whole it appears that whatever discriminating power was present was not sufficient to effectively distinguish

TABLE 4.7A DISCRIMINANT ANALYSIS FOR 1968-1974 CLUSTERS

VARIABLES IN SET 1	STANDARDISED COEFFICIENTS (NO. OF FUNCTIONS=0)
ENERG1	-
ENERG2	-
ENERG3	-
LABQAL	

NO VARIABLES QUALIFIED AS DISCRIMINATORS

VARIABLES IN SET 2 \ FUNCTION NUMBER	STANDARDISED COEFFICIENTS (NUMBER OF FUNCTIONS = 3) 1	2	3
E1RATE	0.60	0.78	0.76
E2RATE	0.61	-1.11	-0.41
E3RATE	0.28	1.35	-0.30
LABQAL	-	-	-
CANONICAL CORRELATION COEFFICIENT	.57	.48	.03
WILKS'S LAMBDA	.52	.77	.99
CHI-SQUARE	18.15	7.36	.02
DEG. OF FREEDOM	9	4	1
LEVEL OF SIGNIFICANCE	.03	.12	.86

PERCENTAGE OF CASES CORRECTLY CLASSIFIED = 43.75%

NOTE: Variables with coefficients marked "-" not selected for inclusion by stepwise procedure.

TABLE 4.7B GROUP MEANS OF DISCRIMINATING VARIABLES

CLUSTER / VARIABLE	HIGH FLYERS	MODERATE GROWTH	STABLE	DECLINING
ENERG1	1.10	1.34	3.28	5.23
ENERG2	0.79	1.29	2.53	6.60
ENERG3	58.39	57.36	50.50	57.15
E1RATE	-0.49	9.93	2.30	-7.57
E2RATE	0.08	9.93	0.50	-3.99
E3RATE	0.87	-0.01	1.67	-4.39
LABQAL	25.72	20.69	30.43	28.46

between the 4 performance groups, because the derived set of classification equations could correctly classify only 43.75 percent of the industries.

The group means for the discriminating variables given in Table 4.7B show that there is a fairly clear pattern of relationships between the discriminating variables and the performance clusters. In particular, the high flyers are characterised by very stable levels of energy consumption, i.e. the rates of change in the variables E1RATE, E2RATE and E3RATE are virtually zero. This is in direct contrast with the discriminant analysis results in Section 3.1, where the high flyers were found to be characterised by relatively high increases in total energy consumption. On the other hand, the declining industries are characterised by decreases in total energy consumption over the time period. The moderate growth industries show definite increases in total energy consumption, and these increases appear to be almost wholly in electricity consumption.

These results suggest numerous possibilities for further research via discriminant analysis. Apart from the obvious possibility of testing other variables and applying discriminant analysis to various pairs of performance groups, the technique could also be applied to identify what characteristics distinguish, say, high flyers in one period from high flyers in another period.

In this way the nature of the relationships between structure and performance can be more fully explored and assessed.

4.8. CLUSTER ANALYSIS ON 2 VARIABLES

Having performed cluster analysis on 32 Tasmanian manufacturing industries at the 3-digit ASIC level for 2 time periods, it was decided to attempt a cluster analysis on the same 32 industries on the basis of 2 variables : performance in the period 1968-1974 and performance in the period 1975-1982. It was expected that such an attempt would ideally result in 4 clusters in which industries are classified as 1) 'good' performance in both periods, 2) 'bad' performance in both periods, 3) 'good' performance in the first period and 'bad' performance in the second, and 4) 'bad' performance in the first period and 'good' performance in the second. Such a result, however, was not to be.

This last cluster analysis resulted in a 4-cluster optimum, as shown in Table 4.8. The industry description for each 3-digit ASIC code is given in Appendix A. The first cluster consisted of 3 industries, the second cluster consisted of 5 industries, the third and largest cluster comprised 16, or half of the total number of industries and the remaining 8 industries comprised the fourth cluster. That this classification is very weak is clearly shown in the individual growth rates

TABLE 4.8 CLUSTER ANALYSIS RESULTS FOR TWO PERIODS

CLUSTER 1	CLUSTER 2	CLUSTER 3	CLUSTER 4
323	217	211	212
348	245	215	213
347	288	216	218
	294	234	235
	323	246	253
		254	295-6
		263	316
		264	335
		275	
		276	
		285	
		286	
		287	
		314	
		315	
		336	

of the industries for each period of time. In the first cluster, the growth rates ranged from -16.92 to 152.34 percent in the first period, and 77.19 to 85.72 percent in the second period. Industries in the second cluster had growth rates ranging from 21.20 to 47.71 percent in the first period and -11.67 to 26.32 percent in the second period. In the third cluster, the growth rates ranged from -9.51 to 14.83 percent and 2.73 to 27.07 percent in the first and second periods respectively. The fourth cluster consisted of industries whose growth rates ranged from -15.04 to 13.33 percent in the first period, and -28.28 to -3.03 percent in the second period.

This unexpected result points to one important consideration in the application of cluster analysis. Care must be exercised when cluster analysis is undertaken on the basis of more than one variable, for the choice of variables, as well as technique is crucial to the structure of the clusters obtained (Everritt, 1974, p.48). It is possible that the chosen variables implicitly impose a certain structure on the clusters to be obtained, so that the true clusters are 'missed' by the clustering algorithm. In other words, the results show that the expected clusters on the 2 variables cannot be obtained from the particular clustering algorithm that was used, i.e. using a Euclidean distance measure to obtain spherical clusters when in fact the clusters are not of a spherical nature. Thus, given this particular clustering algorithm, more meaningful results could perhaps have been obtained if the individual industries had been clustered on different variables over the same time periods rather than the same variable over different time periods. Unfortunately, time constraints have not allowed this alternative to be tested, but it is hoped that better results can be obtained in future research.

CHAPTER 5. SUMMARY AND CONCLUSIONS

5.1 SUMMARY

It was the purpose of this project to attempt to identify any significant relationships which may exist between various structural characteristics of the Tasmanian manufacturing sector and performance as measured by certain indicators, through the use of two statistical techniques, namely cluster analysis and discriminant analysis. Since no major quantitative work has as yet been done for Tasmanian on these structure-performance links, it was hoped that his study would at least provide some useful information which would serve as a basis for further comprehensive research. The original idea for this project was derived from Parry's (1977) paper on "the Structure and Performance of Australian Manufacturing Industries".

Cluster analysis is a method by which individuals are grouped on the basis of one ore more variables, such that each group consists of individuals who are as 'close' as possible on the basis of those variables, yet are as distinct as possible from all other groups. Having obtained these clusters of individuals, discriminant analysis can be applied to determine what other attributes of these individuals distinguish between the clusters. Thus in the context of this study, cluster analysis was used to group Tasmanian manufacturing industries into various performance categories after which discriminant analysis was applied to identify structural characteristics which distinguished between the performance groups.

Although various performance indicators could have been used to cluster the industries, time constraints permitted only the use of one performance indicator, namely growth as measured in terms of the rate of change in an industry's annual contribution to total manufacturing value-added. It was felt that this particular indicator of performance was a reasonable choice since many of the claims concerning structure-performance links in Tasmanian manufacturing reflect a concern for benefits to the State in terms of labour inputs. The structural characteristics variables tested by discriminant analysis were also suggested by various claims made by Callaghan (1976), Wilde (1981), Parry (1977) and Jones (1983). Thus the discriminating variables which were tested fell into four broad categories : resource utilisation in terms of energy consumption and labour skills, protection, export intensity and Tasmanian-based versus mainland-based. It is regretted that currently available data did not permit the inclusion of seemingly important variables such as industry concentration, diversification and transport costs.

Data was constructed at the 3-digit Australian Standard Industry Classification level for Tasmanian manufacturing industry. The main time period under study was 1975 to 1982, but cluster analysis was also applied to the earlier time period of 1968 to 1974 and a comparison made with the original time period.

Cluster analysis on the 32 Tasmanian manufacturing industries included in this study resulted in four performance groups. The were categorised as high flyers, moderate growth industries, stable industries and declining industries. When discriminant analysis was applied to these four groups simultaneously, it was found that significant differences did exist in certain characteristics, mainly in rates of change in energy consumption and in the level of exports interstate, but it should be kept in mind that had a more complete set of variables been available, particularly for export figures, the results may have been slightly different. Nevertheless, the discriminant functions failed to discriminate effectively between the four groups, so that in order to highlight the existing differences it was necessary to resort to applying discriminant analysis to pairs of performance groups.

The best result was obtained on the discriminant analysis between stable industries and declining industries. The function correctly classified 94 percent of the industries and it was found that stable industries were characterised by decreasing total energy consumption with the decrease mainly in electricity usage and that they tended to export a somewhat higher proportion of total exports to overseas destinations.

Since discriminant analysis did not yield very conclusive results on the four performance groups, it was thought that other significant differences may exist between only positive and negative growth categories. However, discriminant analysis on these two groups also failed to identify good discriminating characteristics.

Cluster analysis was also undertaken on 3-digit ASIC level manufacturing industries for the period 1968 to 1974, because the original time period of 1975 to 1982 was a recessionary period and the results obtained could only pertain to such conditions. Here again the industries fell into four performance groups that could be labelled high flyers, moderated growth industries, stable industries and declining industries.

Although data was extremely limited, the rate of change in energy consumption variables emerged as significant discriminators. In particular, it may be noted that the high flyers were characterised by very stable levels of energy consumption, while the high flyers of the 1975-82 period were characterised by very high rates of increase in energy consumption. This would appear to provide some support for the argument that Tasmania needs to increase its energy-generating capacity to meet future demands if this trend continues, but what is not clear is which energy sources should be developed.

Finally, cluster analysis was used to group industries on the basis of two performance variables : growth in the 1968-74 period and growth in the 1975-82 period. It was expected that this would yield clusters of industries in terms of 'good' performance in the first period and 'not-so-good' performance in the second period, and vice versa, etc. However, the results merely yielded a broad grouping of industries into high flyers, moderate growth industries, stable industries and declining industries.

5.2 CONCLUSIONS

The purpose of this project was to provide information on the nature of structure-performance relationships within the Tasmanian manufacturing sector through the use of cluster analysis and discriminant analysis. It has been successful to the extent that these two statistical techniques have proven to be useful tools in the identification of such relationships, and that in the course of the project, some results were obtained which appear to lend some support to some of the existing claims concerning Tasmanian manufacturing, for which there has been a lack of quantitative evidence.

In particular, the results seem to support Wilde's argument that 'filtered-down' industry is characterised by slow expansion, as well as the argument that Tasmania needs to make more provision for future energy requirements. It has not been possible, however, to determine if these energy requirements should be concentrated on the development of water resources for hydro-electricity generation.

Very little can be said, however, on relationships with respect to other characteristics such as export intensity and labour skills. It is clear that more characteristics should be tested in future.

Finally, two important points have to be made concerning the results of this study. The first is that any variables which are identified as discriminators must be considered as such only in relation to all other variables tested at the same time. The discriminant algorithm selects variables on the basis of their contribution to the separation of the populations in question. Therefore a variable which emerges as having some discriminating power when tested with one set of variables may not be selected at all when tested with a different set of variables, if that set contains variables with greater discriminating power. It is therefore important to test as many variables or sets of variables as possible to ensure the validity of discriminant analysis results.

The second point to be made is that in any kind of quantitative work, the results can only be as good as the data input. The potential of this approach to identify structure-performance relationships has been severely constrained by the lack of suitable data on Tasmanian manufacturing, even at the 3-digit ASIC level of disaggregation. As demonstrated in Chapter 3, a considerable amount of time was devoted to obtaining estimates for individual industry value-added, and given the short-term nature of this particular work it was not feasible to attempt the construction of data on all other variables. The consequent dependence on available published data has also rendered the validity of the results subject to the correctness of the published figures.

This study has merely scratched the surface of this potentially rewarding field of research. It is hoped that the work that has been done will serve as a useful starting point for further research.

APPENDIX A. AUSTRALIAN STANDARD INDUSTRY CLASSIFICATION

CODES AND INDUSTRY DESCRIPTIONS

ASIC	INDUSTRY DESCRIPTION
211	MEAT PRODUCTS
212	MILK PRODUCTS
213	FRUIT & VEGETABLE PRODUCTS
214	MARGARINE, OILS & FATS
215	FLOUR MILLS & CEREAL FOOD PRODUCTS
216	BREAD, CAKES & BISCUITS
217	SUGAR & OTHER FOOD PRODUCTS
218	BEVERAGES & MALT
234	TEXTILE FIBRES, YARNS & WOVEN FABRICS
235	OTHER TEXTILE PRODUCTS
244	KNITTING MILLS
245	CLOTHING
246	FOOTWEAR
253	WOOD & WOOD PRODUCTS
254	FURNITURE & MATTRESSES
263	PAPER & PAPER PRODUCTS
264	PRINTING & ALLIED INDUSTRIES
275	BASIC CHEMICALS
276	OTHER CHEMICAL PRODUCTS
285	GLASS & GLASS PRODUCTS
286	CLAY PRODUCTS & REFRACTORIES
287	CEMENT & CONCRETE PRODUCTS
288	OTHER NON-METALLIC MINERAL PRODUCTS
294	BASIC IRON & STEEL
295*	BASIC NON-FERROUS METALS
296*	NON-FERROUS METAL BASIC PRODUCTS
314	FABRICATED METAL PRODUCTS
315	SHEET METAL PRODUCTS
316	OTHER FABRICATED METAL PRODUCTS
323	MOTOR VEHICLES & PARTS
324	OTHER TRANSPORT EQUIPMENT
334	PHOTOGRAPHIC, PROFESSIONAL & SCIENTIFIC EQUIPMENT
335	APPLIANCES & ELECTRICAL EQUIPMENT
336	INDUSTRIAL MACHINERY & EQUIPMENT
345	LEATHER & LEATHER PRODUCTS
346	RUBBER PRODUCTS
347	PLASTIC & RELATED PRODUCTS
348	OTHER MANUFACTURING

* These two industries were grouped as one 3-digit class due to the data problems associated with their previous classification as 292, prior to 1974.

APPENDIX B. DISCRIMINANT ANALYSIS : A GEOMETRICAL PRESENTATION

Bolch & Huang (1974, pp.231-233) present a more sophisticated geometric treatment of discriminant analysis than that given in Chapter 3.

As in the Cooley & Lohnes (1971, p.245) presentation, let there be two populations, I and II, and two variables X_1 and X_2 on which observations can be taken on each member of the populations. Thus the ellipses I and II in Fig. 1 represent the bivariate distributions of populations I and II in standardised form The discriminant problem is to find some linear combination of X_1 and X_2 which will separate the groups as much as possible, i.e. to find a linear combination such that the overlap between the distributions for the two populations is minimised. Note that the closer the means vectors of the two populations, the greater the overlap and the more difficult it becomes to discriminate between the two groups. The linear discriminant function can be expressed as

$$Y_{it} = BX_{i1t} + BX_{i2t} \qquad \text{for } i = 1,2 \quad t = 1...n$$

where Y = discriminant score

B = standardised coefficients

X = discriminating variables

i = population

t = (t)th member of the (i)th population

Bolch & Huang (1974, p.282) stress that "...unlike regression analysis, the variable Y is a result of combining the X variables -- it is not a set of values to be fit by use of the X variables".

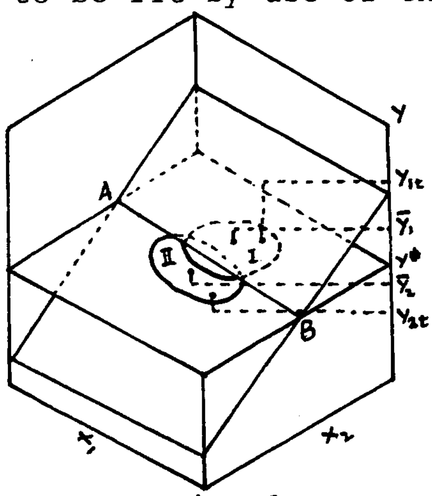


Fig. 1

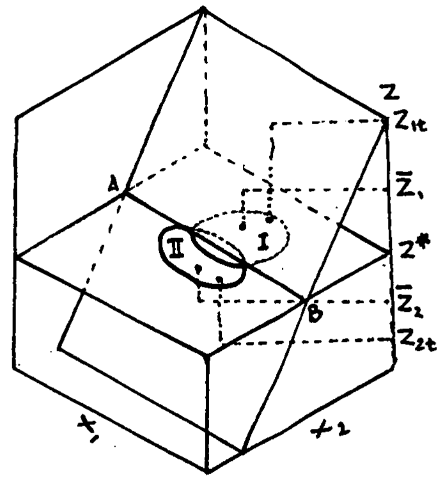


Fig. 2

In geometric terms, the discriminant function defines a plane such that the projection of the two-dimensional scores X_{11t} and X_{12t} on the plane can be transformed into a one-dimensional score Y_{1t} along the Y-axis, as shown in Fig.1. Similarly, the scores X_{21t} and X_{22t} can also be transformed into a one-dimensional score Y_{2t} along the Y-axis. The plane cuts the ellipses at line AB, which can be projected on the Y-axis as Y*. Thus the plane cuts the ellipses such that most of ellipse I lies below the plane while most of ellipse II lies above the plane. Any individual whose scores on X_1 and X_2 are projected on to the Y-axis above Y* will be classified as belonging to population I and any individual whose scores on X_1 and X_2 are projected on to the Y-axis below Y* will be classified as belonging to population II. Thus the one-dimensional score Y* represents the critical discriminant score, d*, as previously discussed in Chapter 3.

Note that misclassification occurs when a member of population I projects below Y* or when a member of population II projects above Y*. The overlap between the two population distributions identifies the area where misclassification can occur, so it is therefore desirable to minimise the overlap, particularly where the cost of misclassification is high.

Since there are an infinite number of discriminant planes passing through the line AB, the problem arises of choosing an appropriate plane for discrimination. The problem becomes clear when we consider Fig. 2, which is similar to Fig. 1 on all but two points :

1) the two-dimensional scores X_1 and X_2 are projected on to the Z-axis so that the discriminant plane is described by

$$Z_{it} = B_1 X_{i1t} + B_2 X_{i2t} \qquad i = 1,2 \quad t = 1,2,\ldots n$$

2) the discriminant plane in Fig. 2 is steeper than that in Fig. 1

If we let the squared distances $(\overline{Y}_1 - \overline{Y}_2)^2$ and $(\overline{Z}_1 - \overline{Z}_2)^2$ represent the amount of separation between the two distributions in Fig. 1 and Fig. 2 respectively, we can see that the steeper discriminant plane in Fig. 2 is superior to that of Fig. 1 because it results in a greater separation of the two populations, i.e. $(\bar{Z}_1 - \bar{Z}_2)^2$ is greater than $(\bar{Y}_1 - \bar{Y}_2)^2$. These squared distances define the 'between-group' variation. At the same time, the squared distance between any projected score Y_{it} and the projected mean Y_i, i.e. $(\bar{Y}_{it} - \bar{Y}_{it})^2$, is less than the corresponding squared distance $(Z_{i\bar{t}} - Z_{i\bar{t}})^2$.These squared distances define the 'within-group' variation. Since large 'within-group' variation is undesirable for statistical analysis the discriminant plane in Fig. 1 is superior to that in Fig. 2. The problem therefore, is to find an optimal discriminant plane that will satisfactorily separate the populations, and yet retain an acceptable degree of within-group variation. One solution to this is to maximize the ratio

$$L = \frac{\text{between-group variation}}{\text{within-group variation}}$$

Although other solutions to this optimisation problem exist, when the population size is small, as is the case for this study, the results obtained are virtually identical. Therefore it was decided to used the criterion of maximising Rao's V, a generalised measure of distance, so as to obtain the greatest overall separation of the populations.

BIBLIOGRAPHY

Anderburg, M.R. Cluster Analysis for Applications, Academic Press, Inc., New York, 1973.

Anderson, T.W. Introduction to Multivariate Statistical Analysis, John Wiley & Sons, New York, 1958.

Australian Bureau of Statistics. Economic Censuses : Manufacturing Establishments, 1968-69, 1970-71, 1972-73 to 1981-82, (Cat. No. 8202.6).

Australian Bureau of Statistics. Tasmanian Yearbook 1980, Government Printer, South Australia, 1980, (Cat. No. 1301.6).

Australian Bureau of Statistics. Tasmanian Yearbook 1983, Government Printer, South Australia, 1983, (Cat. No. 1301.6).

Australian Bureau of Statistics. Tasmanian Yearbook 1984, Government Printer, South Australia, 1984, (Cat. No. 1301.6).

Australian Bureau of Statistics. Trade & Shipping Statistics Tasmania, 1968-69 to 1978-79, (Cat. No. 5401.6).

Bolch, B,W, & Huang, C.J. Multivariate statistical Methods for Business & Economics, Prentice-Hall, Inc., New Jersey, 1974.

Conlon, R.M. "A Discriminant Analysis of Australian and Canadian Manufacturing Industries", Working Paper No. 52, 12th Conference of Economists, Hobart,28 August - 1 September, 1983.

Callaghan, B.E. Inquiry into the Structure of Industry and the Employment Situation in Tasmania, Australian Government Publishing Service, Canberra, 1976.

Caves, R., Ward,I., Williams, W. & Wright, C. Australian Industry : Structure, Conduct, Performance, Prentice-Hall of Australia Pty. Ltd., Sydney, 1981.

Champion, D.J. Basic Statistics for Social Research, Macmillan Publishing Co., Inc., New York, 1981.

Cooley, W.W. & Lohnes, P.R. Multivariate Data Analysis, John Wiley & Sons, Inc., Toronto, 1971.

Duran, B.S. & Odell, P.L. Cluster Analysis : A Survey, Springer-Verlag, Heidelberg, 1974.

Eisenbeis, R.A. "Pitfalls in the Application of Discriminant Analysis in Business, Finance & Economics", Journal of Finance, Vol. 32, No. 3, June 1977, pp.875-900.

Everitt, B.S. Cluster Analysis, Heinemann Educational Books Ltd., London, 1974.

Everitt, B.S. & Dunn, G. Advanced Methods of Data Exploration and Modelling, Heinemann Educational Books Ltd., London, 1983.

Frank, R.E., Massy, W.F. & Morrison, D.G. "Bias in Multiple Discriminant Analysis", Journal of Marketing Research, Vol.2, August 1965, pp.250-258.

Harris, R.J. A Primer of Multivariate Statistics, Academic Press, Inc., New York, 1975.

Industries Assistance Commission. Trends in the Structure of Assistance to Manufacturing, (Information Paper No. 1), Australian Government Publishing Service, Canberra, 1980.

Jones, R."Energy Policy and Tasmania's Electricity Needs".Energy in Tasmania 1980 - 2000,(Public Symposium), ANZAAS Inc., Tasmanian Division, 1980, pp.93-108.

Klecka, W.R."Discriminant Analysis". In Nie, N.H., Bent, D.H. & Hull, C.H.,SPSS : Statistical Packagefor the Social Sciences Manual, 2nd ed., 1975.

Klecka, W.R. Discriminant Analysis, Sage Publications Inc., Beverly Hills, 1980.

Kshirsagar, A.M. Multivariate Anlysis,2nd ed., Marcel Dekker,New York, 1972.

Lorr, M. Cluster Analysis for Social Scientists, Jossey-Bass Inc., San Francisco, 1983.

Morrison, D.G. "On the Interpretation of Discriminant Analysis", Journal of Marketing Research, Vol. 6, May 1969, pp.956-963.

Parry, T.G. "Structure & Performance of Australian Manufacturing Industries", University of New South Wales, 1977.

Robertson, T.S. & Kennedy, J.N. " Prediction of Consumer Innovators : Application of Multiple Discriminant Analysis", Journal of Marketing Research, Vol. 5, February 1968.

Sherman, R. The Economics of Industry, Little, Brown and Company, Inc., Boston, 1974.

Srivastava, M.S. & Carter, E.M. An Introduction to Applied Multivariate Statistics, Elsevier Science Publishing Co., Inc., New York, 1983.

Wilde, P.D. "Industrial Change, Structural Adjustment & the Periphery : The Case of Tasmania". In Linge,G.J.R. & McKay, J. (Eds.). Structural Change in Australia : Some Spatial & Organisational Responses, Research School of Pacific Studies, Australian National University, 1981, pp. 219-35.

Wishart, D. CLUSTAN User Manual, 3rd ed., Program Library Unit, Edinburgh University, 1978.