Recreation Demand Modelling

Occasional Paper

During 1976 and 1977 the Victorian Department of Youth, Sport and Recreation and the Geelong Regional Planning Authority undertook a major study of recreational activity in the Geelong area. As part of that study, John Paterson Urban Systems was commissioned to undertake an extensive household survey and to develop a set of recreation site usage models for locations in the Geelong region. These data, together with information on environmental usage limits for each site in the region were used in the development of recreational area management programmes.

BUREAU OF TRANSPORT ECONOMICS

RECREATION DEMAND MODELLING

John Paterson Urban Systems Pty Ltd

AUSTRALIAN GOVERNMENT PUBLISHING SERVICE
CANBERRA 1978

Printed by Canberra Reprographic Printers, 119 Wollongong Street, Fyshwick, A.C.T.

FOREWORD

During 1976 and 1977 the Victorian Department of Youth, Sport and Recreation and the Geelong Regional Planning Authority undertook a major study of recreational activity in the Geelong area. As part of that study, John Paterson Urban Systems was commissioned to undertake an extensive household survey and to develop a set of recreation site usage models for locations in the Geelong region. These data, together with information on environmental usage limits for each site in the region were used in the development of recreational area management programmes.

The Commonwealth Bureau of Roads was given access to the data collected during the course of the Geelong Recreational Study. This provided the opportunity to review and to further develop the recreational travel demand models developed in that study. The Bureau commissioned John Paterson Urban Systems to assist it in this task and this is a report of their work.

The work was carried out by Mr R. Skinner and Dr J. Symons of John Paterson Urban Systems under the supervision of Mr R. Lombardo of the Bureau. The Bureau would like to thank the Victorian Department of Youth, Sport and Recreation and the Geelong Regional Commission, formerly the Geelong Regional Planning Authority, for allowing the Bureau to use the data collected in the Geelong Recreational Study.

(G. R. CARR)
Assistant Director (Acting)
Transport Planning

Bureau of Transport Economics
Melbourne
November 1978

CONTENTS

Page

TABLES

FIGURES

INTRODUCTION

In January 1976 the Victorian Department of Youth Sport and Recreation in conjunction with the Australian Department of Environment, Housing and Community Development asked John Paterson Urban Systems to undertake a study of the patterns of participation in outdoor recreation at sites within the Geelong Region. Eighteen months later the consultants produced a report[(1)] which describes in detail the models which were developed to predict demand for recreation in the Region.

From the outset the study was seen as innovative in its approach and ambitious in scope, but until now there has been scant opportunity for any critical appriasal of the study procedures. This report examines the models developed in the Geelong Recreation study and discusses the assumptions which underly the structure of the models.

Perhaps the more interesting aspects of the report are the sections which look at ways in which the Geelong Recreation Study approach can be improved upon. In particular, it is suggested that factor analysis should be used as a standard data analysis procedure prior to any modelling. Examining "factors" in the data is seen as advantageous in two ways; firstly, it enables the analyst to specify models with independent variables which are uncorrelated and, secondly, it provides a more comprehensive "feel" for the factors (groups of attributes) which are dominant in influencing behaviour.

This report contains 4 main chapters. Chapter 1 presents a review of the Geelong Recreation Study models. The first part of this chapter is general in its approach and provides the reader with an overview of the models and study findings. The

---

(1) R. Skinner, J. Symons & J. Paterson (1977), The Geelong Recreation Study - Phase 1, Department of Youth Sport and Recreation, Victoria.

second part of this chapter is for readers concerned with a formal description of the models and their assumptions. The final part of the chapter discusses the data on which the validity of the whole modelling exercise rests.

Chapter 2 discusses the need for attitudinal data to increase the behavioural significance of the modelling process. A theoretical framework is presented for the incorporation of attitudinal information into the disaggregate recreation modelling process. The final part of this chapter is an annotated bibliography of some of the more significant work concerned with attitudinal analysis which is directly relevant to the type of demand modelling being discussed in this report.

Chapters 3 and 4 present a discussion of factor analysis and its applications in recreation demand analysis. Chapter 3 introduces the concept of factor analysis and gives an example of its use as applied to the Geelong Recreation Study data. In Chapter 4 the usefulness of factor analysis in building models of recreation demand is discussed further and illustrated by applying the methods to a model of participation in surfboard riding.

CHAPTER 1 - THE MODELS DEVELOPED IN THE GEELONG RECREATION STUDY

The most important requirement of all demand models developed in the study was that the analysis should be sensitive to the factors which influence policy programming. That is, if realistic resource management schemes were to represent the ultimate requirement of the study, then any models developed for use in predicting future area usages should be responsive to future population growth and social trends, and, if possible, the area management schemes themselves.

Following a review of existing recreation demand modelling techniques, it became apparent that most of the demand models estimated to date had excluded the variables which are crucial from a policy point of view. The studies relied on data collected on site which inevitably excludes all members of the population making alternative choices. That is, if one is interested in why people visit the You Yangs(1), for example, interviewing on site provides an indication of present visitor patterns, but throws no light on how competing recreation sites (the Dandenongs(1), King Lake(1) etc) may attract those who decide not to visit the You Yangs. Without information about individuals visiting competing sites it is impossible to formulate predictions concerning future levels of usage under different patterns of population growth and site accessibility.

There are other problems associated with on-site interviews which are hard to overcome in practice. The most significant of these is the problem of controlling the sample frame. Numerous studies can be cited where on-site interviews have been conducted and conclusions reached which are related to the sample, but unrelated to the population of visitors at the beach or park. It is difficult to draw any rigorous conclusions from sample data unless the sample fraction has been accurately assessed. As

---

(1) Recreation parks in Victoria, Australia.

anyone who has attempted the exercise will testify to, counting the number of people in a national park or at a beach on any one day is not always a feasible task.

To overcome the very restrictive problems associated with on-site data collection, it was necessary to interview individuals at their places of residence. A sample of 1787 individuals were interviewed from households drawn at random from the urban areas of Melbourne, Ballarat and Geelong(1). By so doing many individuals who chose not to visit the sites under study were also interviewed.

THE MODEL STRUCTURE - GENERAL DESCRIPTION

The models developed in this study could be described as "disaggregate" - meaning that individual behaviour is observed (via a household questionnaire) and related, in a mathematical way, to certain characteristics of that individual. The logical alternative to a "disaggregate" modelling approach would be "aggregate" modelling. Richards and Ben-Akiva (1975)(2) were amongst the first to argue strongly in favour of disaggregate modelling. They argue that for planning purposes we are concerned with the prediction of the behaviour of aggregates of people. In order to predict the number of people visiting a particular recreation site, we need a demand function that represents the aggregate demand of all potential recreators. Aggregate demand, however, is nothing more than the sum of the demand functions of individual consumers.

In traditional transportation and recreation planning studies data are collected at the level of the individual consumer. In

---

(1) Melbourne is a large conurbation with a population of 2.5 million, 70 kms north of Geelong. Ballarat is the only other major urban centre within 100 kms of Geelong; it has a population of 60,000. Geelong itself has a population of 110,000 and is the central city servicing the Geelong Region. See Figure 1.1

(2) M.G. Richards and M.E. Ben-Akiva (1975), A Disaggregate Travel Demand Model, Saxon House, England.

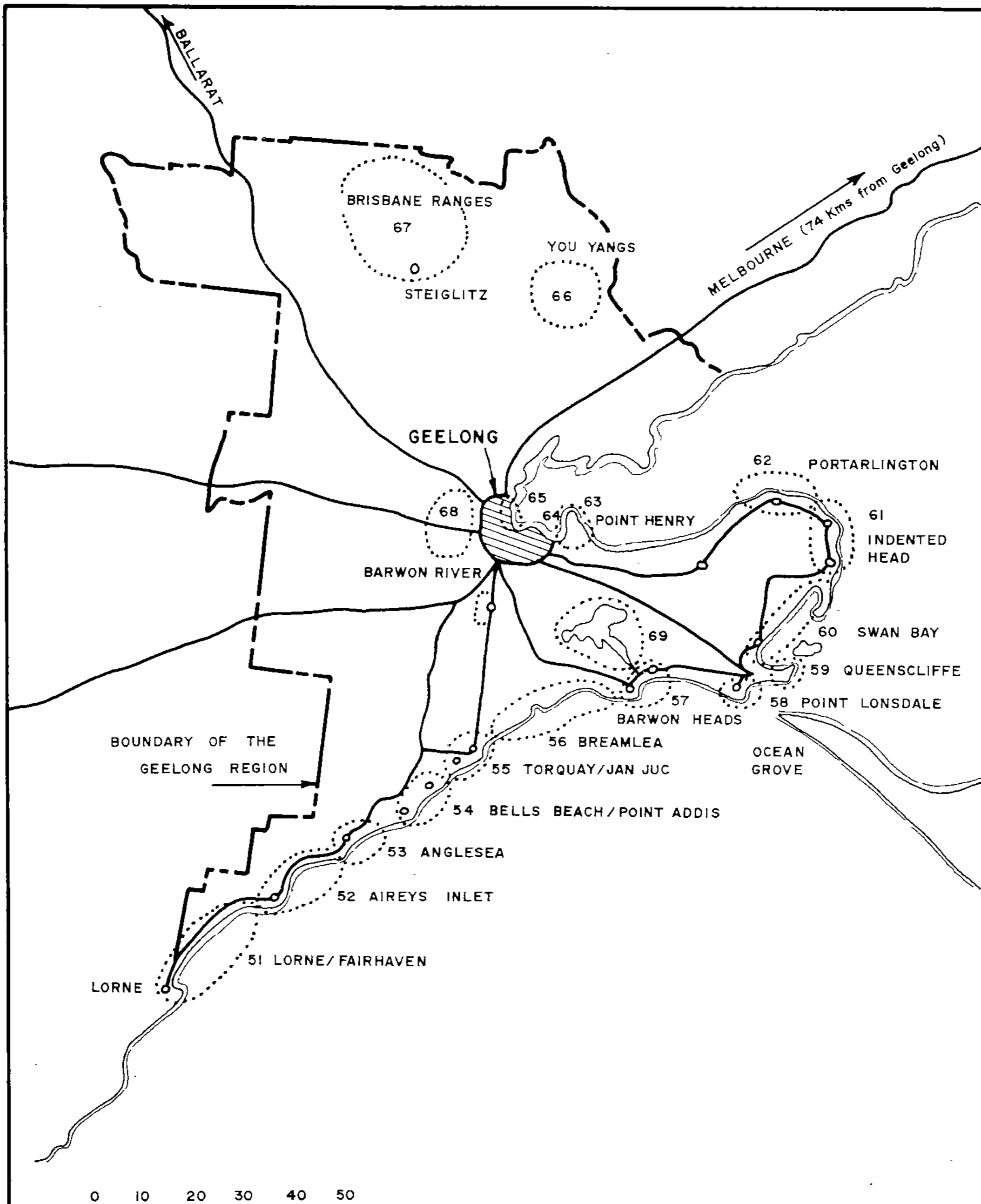


FIG. 1.1 - THE STUDY REGION

aggregate modelling studies the data are then aggregated geographically to zones, demand models are then estimated using the aggregate, or averaged, data, with the zone as the observation unit.

Richards and Ben-Akiva argue that:

> Aggregation before the model construction phase of the analysis will cloud the underlying behavioural relationships and will result in a loss of information. An aggregate model which is based on averages of observations of socio-economic types and geographic location does not necessarily represent an individual consumer's behaviour, nor the average behaviour of the group under a variety of conditions. Thus there is no reason to expect that the same relationship would hold in another instance or another location. Models estimated directly from individual observations, without aggregation, represent the typical behaviour of individual consumers.

At best, therefore, traditional aggregate modelling explains only differences between the zones. Thus zonal aggregation tends to obscure any relationships that might exist between observed travel behaviour and socio-economic characteristics of an individual.

It was for these reasons that "disaggregate" models of recreation demand were developed in the Geelong Study. The remainder of this section describes, in a general way, the overall structure of the models. The next section provides a review of the results of the application of Geelong recreation models. The final part of this chapter presents a formal description of the models specifications.

The rationale underlying the structure of the models used in the Geelong study assumes that an individual's decision to visit a particular site to undertake a recreation activity is a 4 stage sequential choice process. That is, it is postulated that an activity decision can be modelled as a series of 4 interdependent choices. These choices are expressed in mathematical form below:

$X(i)$ - probability that an individual will undertake activity i;

$N(i)$ - the expected frequency with which activity i is undertaken per year on day trips;

$R_i(g)$ - probability that a person will choose region g to undertake activity i for a day trip;

$D_{ig}(k)$ - probability that a person who has chosen the Geelong region (g=1) to undertake activity i, will choose Geelong activity site k, for a day trip.

For any individual therefore, the number of times he will make day trips visits to Geelong site k to undertake activity i (T_{ik}), is assumed to be a multiplicative form of these expressions, such that

$$T_{ik} = X(i) \, . \, N(i) \, . \, R_i(g) \, . \, D_{ig}(k). \qquad \ldots (1.1)$$

This expression does not, however, represent the total expected number of visits of an individual as it does not include recreation trips made whilst on holiday. It is assumed[(1)] that the frequency with which an individual undertakes recreation activities is different when on holidays, and is denoted by $N'(i)$. The data indicated that site choice for activity participation whilst on holidays is predominantly the site at which people are holidaying. That is, individuals tend to throw away their car keys and undertake most of their activities at their chosen holiday site. It was assumed, therefore, that all activities were undertaken at

(1) This theory rests on a large number of other assumptions which are discussed more fully in the section entitled Formal Presentation of the Models.

the holiday site. To complete the modelling process it is necessary to analyse the likelihood of an individual choosing Geelong site m to holiday in. This is done by defining, firstly:

A(g) - proability of a person choosing the Geelong Region g to holiday in, and secondly,

$S_g(m)$ - probability that a person choosing to holiday in the Geelong region (g=1) will choose to holiday at site m.

When the holiday site choice models are entered into the calculations, the 4-step day trips model is expanded such that the number of visits an individual will make to site k to undertake activity i in any year whilst on holiday is:

$$T'_k = A(g) \cdot S_g(m) \cdot X(i) \cdot N'(i) \qquad \ldots (1.2)$$

The total number of trips to site k by individual i is, therefore, the sum $T_{ik}+T'_{ik}$.

The foregoing is placed into context diagramatically by Figure 1.2. The figure is drawn to represent the choice process which is assumed to lead an individual to choose a particular Geelong site to undertake a particular activity on a given number of times a year.

The models involving probabilities have been calculated using a random utility model. It is postulated that if an individual is faced with a choice between two alternatives he will choose the alternative that gives him the greater utility. That is, if individual t has to choose between alternatives a and b he will choose alternative a if $U_{at} > U_{bt}$, where U_{at} represents utility of individual t gained from adopting alternative a.

It is assumed that $U_{at} = U(Z_a, S_t)$, where Z_a is a vector of attributes of alternative a, S_t is a vector of socio-economic character-

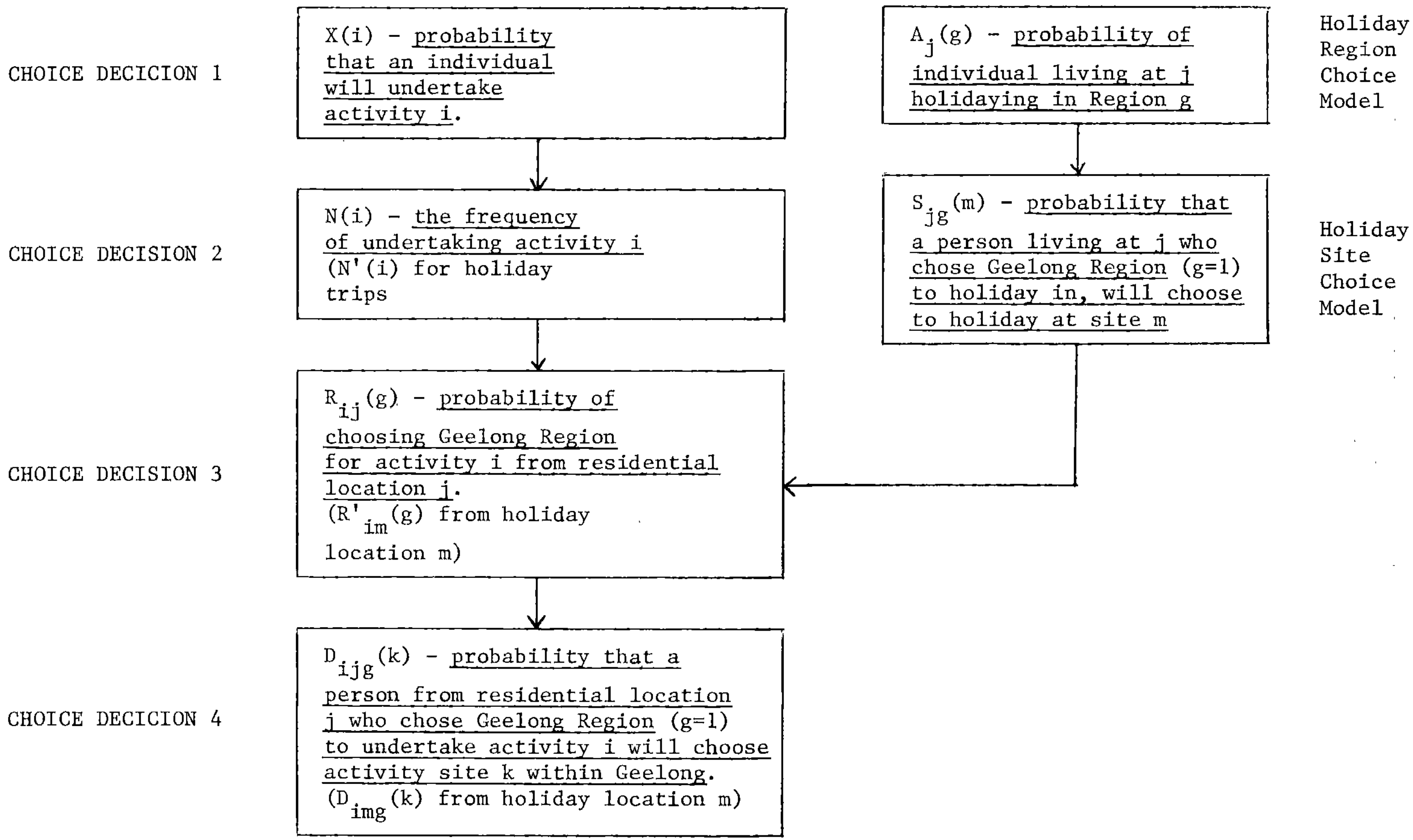


FIG. 1.2 - THE CHOICE SEQUENCE UNDERLYING MODEL DEVELOPMENT

istics of individual t, and $U(Z_a,S_t)$ is the utility function that applies to all alternatives and all consumers.

By observing (via the household interview survey) the actual choices and the choice sets of individuals (or their "revealed preferences"), estimates of the parameters of the utility function U_{at} can be obtained.

It has been assumed that the probability of an individual choosing alternative a from a full set of alternatives A_t can be expressed using a multi-nomial logit model of the form:

$$P(a:A_t) = \frac{e^{U(Z_a,\ S_t)}}{\sum_{a=1}^{g} e^{U(Z_a,S_t)}} \qquad \ldots (1.3)$$

where g = number of alternatives.

The coefficients of the utility function are estimated from observations of actual choice using a statistical maximum likelihood package.

GENERAL REVIEW OF RESULTS OF GEELONG RECREATION STUDY

In the previous section we discussed in broad terms the structure of the models. In the Geelong study the activities (I) for which models were developed were:

- Visiting the beach
- Surf swimming
- Protected Beach Swimming
- Surfboard Riding
- Fishing
- Bushwalking
- Picnicking and barbecuing
- Sight Seeing Driving

Whilst the household survey collected information on 23 specific recreation sites within the Geelong Region, total usage estimate models were formulated for the following sites within the Geelong Region (refer Figure 1 for general location).

51. Lorne
53. Anglesea
54. Bells Beach
55. Torquay
56. Breamlea
57. Barwon Heads/Ocean Grove
58. Point Lonsdale
59. Queenscliff
61. Indented Heads/St Leonards
62. Portarlington
64. Easter Beach
66. You Yangs
67. Brisbane Ranges

In the following six sections we present a summary description of each of the models developed to predict individual behaviour corresponding to each of the 6 choice decisions as discussed earlier.

<u>The Decision to Participate in Activity at Least Once</u> (CHOICE DECISION ONE)

Table 1.1 presents a summary of those factors found to have a significant effect on the probability that an individual will undertake an activity.

A positive influence is one which raises individual probability of participation. That is, the table indicates how each variable influenced the magnitude of the utility function of individuals when choosing each particular activity. The utility function is that described as $U(z_a, S_t)$ in equation 1.3. The logit-based analysis is probabilistic rather than deterministic in that the

TABLE 1.1 - THE EFFECT WHICH EACH VARIABLE HAS ON THE PROBABILITY OF UNDERTAKING EACH ACTIVITY

	Activities							
	Visiting Beach	Surf Swimming	Protected Beach Swimming	Surf Board Riding	Fishing	Bushwalking	Picnicking	Sight Seeing Driving
If MALE	-	+		+	+		-	
If Completed Secondary Education	+	+	+			+	+	
If Employed in Passive Employment	+	+	+	+			+	
Distance to Nearest Facility	-	-		-		-	-	-
If aged 13-20		+		+				
If aged 21-34		+	+	+		-	+	
If aged 35-60			-			-		
If aged > 60 years	-						-	-
Adjusted Income					+	+		
Born in Australia	-	-	-	+				
No. of Children		+	+		+	+	+	+
No. of Babies				-				
No Access to Car	-	-	-	-	-	-	-	-
If Unemployed		+		+				

Blank cells indicate that the particular characteristics did not have a statistically significant effect on the individual's probability of undertaking the activity in question, or that the variable was not tested for statistical reasons (see page 59).
+ = variable has a positive influence on probability of participating in activity
- = variable has a negative influence on probability of participating in activity.

final analysis is of the form "The probability of an individual with particular socio-economic characteristics undertaking a particular activity is x" rather than "An individual with particular socio-economic characteristics will or will not undertake a particular activity".

The probabilities associated with an individual with particular socio-economic characteristics undertaking each of the eight activities were estimated in turn. These models of activity participation have been applied to residents in each of 29 localities of residence to predict the number of individuals undertaking activities at least once in 12 months. These results are presented in Table 1.2. The reader is referred to the Geelong Recreation Study Report(1) for details of form of the models.

The Frequency with which Activities Will be Undertaken (CHOICE DECISION TWO)

Table 1.3 presents a summary of those factors which have a significant influence on the frequency with which an individual undertakes a particular activity, given he undertakes the activity at least once.

It should be remembered that these results are relevant only for individuals who have undertaken the activity at least once, which accounts for the occasional result that may appear to be counter intuitive.

The frequency analysis was undertaken for activities undertaken whilst on holidays as well as for home based trips. Table 1.4 presents the results of the home based frequency models applied over all individuals who undertook an activity at least once (those individuals in Table 1.2). Table 1.5 is for holiday trips. Tables 1.4 and 1.5 represent the total number of activity trips being made from each residential zone.

(1) Skinner et.al., op.cit.

Factors which Affect Region Choice for Undertaking Activities
(CHOICE DECISION 3)

A number of model structures was tested to assess the factors which have most influence on regional choice behaviour. Given that an individual undertakes a particular activity we are interested to know what factors influence the choice of region in which he will do so. Specifically it is sought to discover what influences people to choose the Geelong region in preference to, say, the Mornington Peninsula.

Three regional choice models were estimated. The first for beach activities defined as (1) visiting beach, (3) protected beach swimming, (5) fishing, and (10) power boating. If an individual undertook any of these activities the 'beach activities regional choice' model produces an estimate of the probability of choice of the Geelong region in which to perform the activities concerned. Socio-economic factors which were demonstrated to have a significant effect on region choice were an individual's adjusted income and occupation (blue collar or otherwise).

It turned out that dominant influence on region choice behaviour for beach activities was the generalised cost of travel from an individual's place of residence to the site in question. "Generalised cost" is a concept borrowed from transport planning; it is a linear combination of the vehicle cost of travel and a "cost" associated with the time taken to make the trip. Travel costs influence Regional Choice in different ways for different activities.

The two other regional choice activity groupings were for surf-beach activities (surf swimming and surboard riding) and non-beach activities (bushwalking, picnicking and sight seeing driving).

It is interesting to note that socio-economic factors were not seen to be significant in regional choice decisions for these two activity groups with generalised cost of travel being the dominant factor.

TABLE 1.2 - THE NUMBER OF INDIVIDUALS FROM EACH RESIDENTIAL ZONE PREDICTED TO HAVE UNDERTAKEN ACTIVITIES AT LEAST ONCE IN PAST 12 MONTHS

Residential Zone	Visiting Beach	Surf Swimming	Protected Beach Swimming	Surf Board Riding	Fishing	Bushwalking	Picnicking	Sight Seeing Driving
Melbourne, Sth Melbourne, Port Melbourne	42732	21171	25642	2691	17446	29655	55618	60225
Fitzroy, Collingwood, Richmond	25138	10726	15079	1394	9214	15263	31903	37920
Prahran, St Kilda, Malvern, Caulfield, Brighton	125554	68283	75421	8489	52723	95457	164256	168837
Kew, Hawthorn, Camberwell	71702	40136	43326	4816	32508	60373	97219	97928
Oak.Moorbn.Sndhm.SpringVa.Dand. Mordlic.Chlsea.Frkstn	224284	132649	135167	23727	106409	180864	308948	336089
Box Hill, Nunawading, Ringwood, Waverley, Croydon, Knox	187510	112804	116809	14857	94131	172560	273080	285203
Doncaster, Eltham, Diamond Valley	76920	44541	48863	5178	40021	65375	112772	121440
Coburg, Preston, Heidelberg	87119	45791	53566	5640	41215	61088	118123	134966
Essendon, Brunswick, Northcote	55773	31587	40285	3766	28241	43402	86075	99015
Footscray, Sunshine	56134	29255	34629	3852	26124	40270	75681	87132
Keilor	26510	15046	16482	1882	13022	19870	36382	41264
Williamstown	11176	5860	6731	917	4993	7483	14577	16719
Broadmeadows	46743	26130	29370	2894	23240	32046	63002	73406
Melton	4654	3017	3003	458	2694	4080	6962	7730
Werribee	12141	8284	7758	1421	6725	9530	16721	18370
Ballarat	15260	10210	12315	1262	8920	11894	25914	31642
Urban Geelong/Bellarine	15516	10419	9510	2378	8423	12596	21518	24080
Geelong West	6560	3451	3737	884	2542	4363	8671	10113
Newton	4644	2958	2876	550	2115	3411	6218	7041
South Barwon	12779	8553	7582	2020	6272	9299	16920	19019
Angelsea	953	525	576	156	463	933	1463	1555
Torquay	911	625	542	177	434	688	1216	1390
Barwon Heads/Ocean Grove	1915	1199	1108	359	890	1341	2421	2832
Pt. Lonsdale, Queenscliff	1139	725	671	200	572	701	1410	1703
Portarlington/St Leonards	860	538	542	135	431	504	1045	1349
Leopold	1381	945	870	288	647	965	1887	2162
Corio	20038	13104	12432	3038	9571	16595	28081	31323
Altona	12161	6942	7349	1061	5578	7876	15496	18135

TABLE 1.3 - THE EFFECT WHICH EACH VARIABLE HAS ON THE FREQUENCY WITH WHICH EACH ACTIVITY IS UNDERTAKEN

	Activities							
	Visiting Beach	Surf Swimming	Protected Beach Swimming	Surf Board Riding	Fishing	Bushwalking	Picnicking	Sight Seeing Driving
If MALE								
If Completed secondary Education		-			+			-
If Employed in Passive Employment		+						
Distance to Nearest Facility	-	-	-	-			-	
If aged 13-20		-	+		-			
If aged 21-34								
If aged 35-60								
If aged > 60 years							-	+
Adjusted Income		-	-					
Born in Australia	+	-			-		+	
No. of Children		-	-		-	+	+	
No. of Babies			-	+				+
No Access to Car	+	+			-		-	-
If Unemployed	+				-		-	+

Blank cells indicate that the particular characteristics did not have a statistically significant effect on the individual's probability of undertaking the activity in question, or that the variable was not tested for statistical reasons (see page 59).
+ = variable has a positive influence on probability of participating in activity.
- = variable has a negative influence on probability of participating in activity.

TABLE 1.4 - TOTAL NUMBER OF ACTIVITY TRIPS ESTIMATED TO BE MADE FROM EACH ZONE IN A YEAR - DAY TRIPS ONLY

Residential Zone	Visiting Beach	Surf Swimming	Protected Beach Swimming	Surf Board Riding	Fishing	Bushwalking	Picnicking	Sight Seeing, Driving
Melbourne, Sth Melbourne, Port Melbourne	230754	78641	185484	7726	92818	157173	184751	277034
Fitzroy, Collingwood, Richmond	134952	51102	129681	4455	47985	78977	106640	171622
Prahran, St Kilda, Malvern, Caulfield, Brighton	645569	242953	499011	29292	312348	473076	548816	805906
Kew, Hawthorn, Camberwell	361500	126861	270951	15184	201824	275490	340265	477728
Oak.Moorbn.Sndhm.SpringVa.Dand. Mordlic.Chlsea.Frkstn	1154350	620063	965991	178761	641239	775971	1091800	1652070
Box Hill, Nunawading, Ringwood, Waverley, Croydon, Knox	881664	367718	731231	57584	556782	617826	1086240	1427850
Doncaster, Eltham, Diamond Valley	342848	119088	296631	13705	228828	266801	417496	609749
Coburg, Preston, Heidelberg	423410	149725	365944	14801	244605	335479	359082	655773
Essendon, Brunswick, Northcote	332507	115966	288959	10780	159421	241294	264387	475270
Footscray, Sunshine	275875	110539	244188	12755	152237	196789	244091	423676
Keilor	127725	51155	112076	6586	76827	89413	120059	206321
Williamstown	59232	26956	49808	4402	29458	41156	43731	80252
Broadmeadows	222463	76048	195825	6534	137645	174287	185507	361268
Melton	19082	9354	16216	2108	14817	17544	23670	40196
Werribee	59489	29821	44996	8099	37659	45745	51836	93689
Ballarat	41201	34714	54186	3532	49951	73744	69969	155046
Urban Geelong/Bellarine	83786	61924	63306	38046	51748	56681	68859	127299
Geelong West	34766	28642	32139	13267	14742	22685	28615	48542
Newton	24612	21593	21855	11225	12478	16374	19899	35204
South Barwon	66453	54741	53830	36368	38889	45567	52452	96998
Angelsea	5637	4493	4348	2969	2379	3249	6522	7399
Torquay	4920	5311	4501	5313	2776	3096	3770	7364
Barwon Heads/Ocean Grove	9956	9709	8972	8257	5962	7509	6778	14160
Pt. Lonsdale, Queenscliff	6264	5146	4968	3999	3721	5189	3244	8683
Portarlington/St Leonards	4543	4030	4530	2032	2906	3525	2330	7012
Leopold	6580	6465	7095	4097	3700	5019	5629	10498
Corio	97818	88614	94243	42113	60162	62711	101649	156615
Altona	63235	31238	55117	5943	33468	40170	46489	90676

TABLE 1.5 - TOTAL NUMBER OF ACTIVITY TRIPS ESTIMATED TO BE MADE BY RESIDENTS OF EACH ZONE IN A YEAR WHILST ON HOLIDAYS

Residential Zone	Visiting Beach	Surf Swimming	Protected Beach Swimming	Surf Board Riding	Fishing	Bushwalking	Picnicking	Sight Seeing Driving
Melbourne, Sth Melbourne, Port Melbourne	308601	147703	202574	12247	81620	128350	320424	380481
Fitzroy, Collingwood, Richmond	178093	80659	122615	5698	46759	67264	176456	221155
Prahran, St Kilda, Malvern, Caulfield, Brighton	898017	457492	565021	41995	242530	401409	948512	1171970
Kew, Hawthorn, Camberwell	512477	268807	315057	24906	151266	247528	571449	712336
Oak.Moorbn.Sndhm.SpringVa.Dand. Mordlic.Chlsea.Frkstn	1470830	1003990	971871	151713	509885	717822	1761540	2222210
Box Hill, Nunawading, Ringwood, Waverley, Croydon, Knox	1256860	801526	785235	85609	439540	753574	1529000	1843990
Doncaster, Eltham, Diamond Valley	514087	302553	310962	27081	187241	250918	620095	745335
Coburg, Preston, Heidelberg	622370	311467	388467	27033	198551	252411	673854	890418
Essendon, Brunswick, Northcote	474941	222015	300659	16728	134028	183778	473193	630028
Footscray, Sunshine	392431	212516	261502	19074	123733	163967	423811	555474
Keilor	170301	109833	117021	10162	61201	77491	200099	251712
Williamstown	77113	43364	51827	4850	23966	31428	83089	110346
Broadmeadows	316026	182634	212813	14281	109758	128182	359810	469560
Melton	31648	21122	17718	2796	11854	15504	36897	47152
Werribee	80129	61299	53530	9946	28245	36215	91967	117570
Ballarat	149544	73512	67733	6055	47275	49956	147711	193017
Urban Geelong/Bellarine	95979	89014	62355	20076	37904	47310	116199	166543
Geelong West	43949	28297	24665	5395	13980	18759	46824	60677
Newton	28791	26918	19267	5150	10363	13645	34201	45061
South Barwon	74120	72703	48523	16770	29481	36267	93060	121722
Angelsea	6781	4860	4232	1205	2240	3677	8108	10068
Torquay	5011	5561	3362	1559	2256	2683	6810	9310
Barwon Heads/Ocean Grove	12062	9949	7864	2836	4627	5632	14523	29674
Pt. Lonsdale, Queenscliff	7289	5943	4498	1639	2862	3015	8322	12769
Portarlington/St Leonards	5711	4544	4308	989	2248	2115	6436	11058
Leopold	8923	7498	6535	2099	3595	4030	11878	14395
Corio	124791	111200	89270	22426	47947	63061	160200	203900
Altona	75396	54145	55852	6262	26216	31506	86780	112437

For each activity the annual number of trips made from each zone has been calculated (see Table 1.4). These trips have been apportioned to the Geelong Region and elsewhere using the three activity group regional choice models. Table 1.6 represents those trips (from Table 1.4) which are predicted to have been undertaken in the Geelong region over the 12 months prior to interview.

Factors which Affect Site Choice for Undertaking Activities (CHOICE DECISION 4)

A number of model structures were tested to assess the factors which had most influence on sites within the Geelong Region for each of the eight activities. In Chapter 5 of the Geelong Recreation Report it is demonstrated that generalised cost of travel is the dominant factor for all activities. The magnitude of the effect of generalised cost of access upon choice probabilities varied between activities. It appeared, for example, that changes in generalised cost had least effect on site choice for surfboard riders; differences in cost of travel are presumably regarded as less important than finding good surf. On the other hand, picnickers and bushwalkers seem to be more sensitive to generalised cost of access, closer sites are more readily accepted.

In addition to generalised cost of access a variable which described the physical size of the site was examined. For beach activities a variable related to the length of beach at the site was found to have a positive and significant effect on site choice for all activities except surf swimming. That is, the longer the beach at a site the higher the probability (all other things being equal) that an individual will choose the site.

For non-beach activities a variable related to the area of designated national park was analysed. It was found that such a site characteristics did not have a significant effect on site choice for bushwalking and picnicking, although area of park does positively influence site choice for sight seeing driving.

TABLE 1.6 - ESTIMATED NUMBER OF ACTIVITY DAY TRIPS TO THE GEELONG REGION MADE FROM EACH RESIDENTIAL ZONE

Residential Zone	Activity Group I (a)	Activity Group II (b)	Activity Group III (c)
Melbourne, Sth Melbourne, Port Melbourne	25783	24344	22652
Fitzroy, Collingwood, Richmond	11910	14641	11937
Prahran, St Kilda, Malvern, Caulfield, Brighton	50508	45200	52024
Kew, Hawthorn, Camberwell	27994	25511	25633
Oak.Moorbn.Sndhm.SpringVa.Dand. Mordlic.Chlsea.Frkstn	36824	40588	55464
Box Hill, Nunawading, Ringwood, Waverley, Croydon, Knox	37601	34219	52864
Doncaster, Eltham, Diamond Valley	20329	32761	26872
Coburg, Preston, Heidelberg	43781	47527	42269
Essendon, Brunswick, Northcote	31995	34256	33116
Footscray, Sunshine	53514	61063	42339
Keilor	11975	20383	10898
Williamstown	11315	16329	10978
Broadmeadows	20699	75316	17736
Melton	5337	9028	4747
Werribee	27141	30211	19888
Ballarat	105827	37171	66979
Urban Geelong/Bellarine	194785	99725	212606
Geelong West	79218	41798	83540
Newton	57802	32750	60882
South Barwon	156275	90920	166111
Angelsea	11914	7389	14732
Torquay	12008	10602	12729
Barwon Heads/Ocean Grove	24501	17924	25489
Pt. Lonsdale, Queenscliff	14868	9138	15997
Portarlington/St Leonards	11909	6959	11745
Leopold	17069	10549	18929
Corio	239894	130172	251835
Altona	16333	25113	15425

(a) Activity Group I - Visiting Beach, Protected Beach Swimming, Fishing.
(b) Activity Group II - Surf Swimming and Surf Board Riding.
(c) Activity Group III - Bushwalking, Picnicking, Sight Seeing Driving.

Holiday Site Choice

The holiday site choice process is illustrated in Figure 1.2. The first choice an individual makes when choosing a place to holiday is at the Regional level, designated as $A_j(g)$. A number of socio-economic characteristics were tested in the calibration process, including:

- Blue Collar Occupation
- Adjusted Income
- Age
- The probability of an individual undertaking activity i, i=1, 8, given activity i can be undertaken at that site (REASACT). (This variable was chosen in an attempt to provide some interactive effect which would describe the observed fact that people who are keen to undertake certain activities will choose holiday sites at which that activity can be undertaken.)
- Generalised cost of travel from an individual's residential location to the region in question.

The results of the model calibration indicate that the probability of an individual choosing a particular region is significantly affected by the generalised cost of travel to that region (t=-5.10), but socio-economic variables do not appear to influence the choice process to any significant extent.

The same socio-economic variables were chosen for calibration of the model of specific Geelong site choice selection for overnight stops.

The variable found to be most useful in the holiday site choice model was REASACT. That is, there appears to be a significant connection between an individual who has a high propensity to undertake surfbeach activities (for example) and the probability that the individual may wish to holiday at surf beach locations.

Whereas regional choice was seen to be dependent on generalised cost of travel, it is assumed that, given that an individual chooses to holiday in the Geelong Region, choice of holiday site is independent of travel cost to that site. That is, whilst the decision to go to the Geelong region may be made on generalised cost criteria, the influence of the marginal differences in generalised cost travel to each of the sites within Geelong is not an important determinant.

The Aggregated Models

Table 1.7 presents the total number of trips to each Geelong site by activity type. These estimates have been derived in the following way:

- the day trip component has been computed by multiplying the number of activity trips to the Region (Table 1.6) by the appropriate site choice probabilities as developed in the previous section;
- the holiday trip component has been calculated by multiplying the annual number of activity trips made whilst on holidays (Table 1.5) by the probability of an individual choosing each site to undertake holidays.

Table 1.8 presents the site visit estimates broken down by residential location of the visitors. Estimates of the visits by holidayers and day trippers is presented in Table 1.9.

FORMAL PRESENTATION OF THE MODELS

In this section we shall give a detailed description of the model. The formal presentation given here necessitates a more succinct definition of the choice events than the general treatment given in the earlier sections of this chapter. It is for this reason that the notation of many of the events has been changed although the final models represent exactly the same specification.

TABLE 1.7 - ANNUAL FREQUENCY OF VISITS TO SITES BY ACTIVITY ('00's) PREDICTED FOR JAN-DEC 1976

Sites	Visiting Beach	Surf Swimming	Protected Beach Swimming	Surf Board Riding	Fishing	Bush Walking	Picnicking	Sight Seeing Driving	Total
Lorne/Fairhaven	3710	943	794	174	539	472	206	3004	9842
Aireys Inlet	-	1309	-	140	1431	1857	90	54	4884
Anglesea	1575	6528	363	429	1585	1884	374	661	13404
Bells Beach/Pt Addis	217	328	-	789	351	-	-	31	1766
Torquay/Jan Juc	2492	10991	2095	1016	570	1496	312	994	19967
Breamlea	256	666	-	941	-	-	-	18	1881
Barwon Heads	826	2890	937	312	829	1312	576	1021	8706
Pt Lonsdale	378	1021	208	97	921	97	31	1394	4151
Queenscliff	170	-	1237	-	551	167	167	1038	1540
St Leonards/Indented Hds	41	-	16	-	1661	137	94	239	2187
Portarlington	1584	-	2429	-	347	101	596	864	5923
Eastern Beach/Geelong	1740	-	1808	-	1033	-	202	530	5313
You Yangs	-	-	-	-	-	761	2806	564	3413
Brisbane Ranges	-	-	-	-	-	325	1278	569	2172

TABLE 1.8 - ANNUAL FREQUENCY OF VISITS TO SITES BY RESIDENTIAL ORIGIN OF TRIP MAKER ('00's) - TOTAL FOR HOLIDAY AND HOMEBOUND - JAN-DEC 1976

	Area of Residence								
	East of Melbourne	West of Melbourne	Geelong City	Geelong West	Newtown	South Barwon	Corio	Other Areas	Total
Lorne	4329	3562	168	151	119	329	456	728	9842
Aireys Inlet	2392	1960	46	40	40	116	102	185	4882
Anglesea	5130	4356	348	319	282	802	900	1255	13404
Bells Beach	402	452	88	74	63	199	224	262	1766
Torquay	5436	6286	746	734	636	1806	2025	2293	19967
Breamlea	370	434	90	79	65	202	230	410	1881
Barwon Hds/ Ocean Grove	2600	2331	237	306	282	849	754	1250	8706
Pt Lonsdale	1149	1100	189	161	96	236	469	745	4151
St Leonards/ Indented Hds	528	543	119	86	56	152	274	434	2187
Portarlington	2525	2102	122	113	65	159	324	508	5923
Eastern Beach	1290	1530	211	198	123	298	572	1090	5313
You Yangs	604	1085	181	156	96	235	451	601	3413
Brisbane Ranges	378	648	96	87	55	140	379	387	2172

TABLE 1.9 - ANNUAL FREQUENCY OF VISITS TO SITES BY TYPE OF TRIP PREDICTED FOR JAN-DEC 1976

Sites	From Home	On Holiday	Total
Lorne	2840	7010	9852
Aireys Inlet	524	4359	4884
Anglesea	3289	10114	13404
Bells Beach	1234	532	1766
Torquay	8589	11378	19967
Breamlea	1535	346	1881
Barwon Heads	3536	5170	8706
Pt Lonsdale	1751	2400	4151
Queenscliff	1157	1390	2547
St Leonards	594	1593	2187
Portarlington	1643	3650	5293
Eastern Beach	3956	1356	5313
You Yangs	3413	-	3413
Brisbane Ranges	2172	-	2172

As discussed earlier, there are two models - one for home-based trips and one for holiday-based trips. We shall begin with a discussion of the latter.

The Holiday Model

The basic dependent variable is the number of holiday based activity trips to a given site from another given (holiday) site. The choice mechanism producing these trips is structured as follows:

(M1) the individual chooses to participate in activity i

(M2) the individual chooses a number of overnight trips

(M3) for each overnight trip the individual chooses a region in which to stay

(M4) for each trip directed to a region the individual chooses a site

(M5) for each overnight site choice the individual chooses an activity level

(M6) for each holiday based activity trip the individual chooses a region

(M7) for each holiday based activity trip the individual chooses a site within the region.

The stochastic events corresponding to these choices are set out in Table 1.10.

The alert reader may ask why we have introduced the notion of region into the model - after all a region is merely a collection of sites and choice between regions may be considered to be choice between these sets. There are two reasons for this inclusion - one technical and the other theoretical. In the first place, a model of simultaneous choice between all recreational sites in Victoria (Australia, World?) would demand a huge amount of detailed information and be very difficult to calibrate. In the second place, there is some evidence that individuals perceive a region as an entity in itself - in other words, that there is an interaction between perception of proximate sites to produce an integrated perception of the region as a whole.

The dependent variable in the holiday model is the expected number of annual activity i trips an individual will make from holiday site m in region g to site k in h - $E(T|M_m.K_k.I_i)$, where T is the number of holiday based activity i trips undertaken by the given individual. This quantity is decomposed in the following way (refer to Table 1.11).

$$E(T|M_m.K_k.I_i) = \sum_{s,t} (M1)(s)(M2)(M3)(M4)(t)(M5)(M6)(M7) \quad \ldots (1.4)$$

By repeated application of Bayes' law the right-hand side may be reduced to:

$$\sum_{s,t} stP(I_i.S_s.G_g.M_m.T_t.H_h.K_k)$$

$$= \sum_{s,t} stP(S_s.T_t.M_m.K_k.I_i), \quad \ldots (1.5)$$

which justifies the equality in (1.4).

It should be remembered that (1.4) is not a model but a tautology, and a different set of models, M1 to M7, would be obtained for a different decomposition of $E(T|M_m.K_k.I_i)$ - for example, the roles of S_s and T_t could be reversed. This particular selection was made on the grounds that each model corresponds fairly directly to a genuine recreation decision. If this is so we can hope that the functional dependence of the various probabilities on the socio-economic descriptors of the individual will be simple and readily estimable. Moreover, when we come to make simplifications to the model, these can only be given a behavioural interpretation if a genuine behavioural interpretation can be given to each of the sub-models.

In fact our data showed that there were very few activity trips from a holiday site to a rival site. Accordingly we assumed:

TABLE 1.10 - THE EVENT SPACE

Event	Description
I_i	a given individual participates in activity i
S_s	a given individual undertakes s overnight stays
G_g	a given individual directs an overnight stay to region g
M_m	a given individual directs an overnight stay to site m
T_t	a given overnight stay generates t activity i trips
H_h	a given holiday-based activity i trip is directed to region h
K_k	a given holiday-based trip is directed to site k

TABLE 1.11 - THE HOLIDAY MODEL

Model	Dependent Variable	Description
M1	$P(I_i)$	probability of participation in i
M2	$P(S_s \mid I_i)$	probability of s overnight stays, given participation
M3	$P(G_g \mid I_i.S_s)$	probability that an overnight stay is directed to region g given participation and holiday level s
M4	$P(M_m \mid I_i.S_s.G_g)$	probability that an overnight stay is directed to site m, given participation, holiday level s, and selection of region g
M5	$*P(T_t \mid I_i.S_s.G_g.M_m)$	probability that a given overnight stay generates T_t activity i trips, given that the individual participates etc.
M6	$*P(H_h \mid I_i.S_s.G_g.M_m.T_t)$	probability that a given holiday-based activity i trip is directed to region h given etc.
M7	$*P(K_k \mid I_i.S_s.G_g.M_m.T_t.H_h)$	probability that a given holiday-based activity i trip is directed to site k in h given that it is directed to h etc.

* Note that G_g may be removed from these expressions.

ASSUMPTION 1: Activity trips generated by overnight stays are undertaken only at the site of the overnight stay.

We shall so define our collection of sites so that activity i is possible at each. This definition makes our next assumption fairly reasonable.

ASSUMPTION 2: $P(T_t|I_i.S_s.M_m) = P(T_t|I_i.S_s)$

This says effectively that once the activity is specified, given overnight stays generate the same number of activity trips. It must be borne in mind that our collection of sites is so defined that activity i is possible at each. Thus for example, given an overnight stay, each surf site is expected to produce the same number of surf trips.

Assumption 1 reduces (1.4) to

$$E(T|M_m.K_k.I_i) = \sum_{s,t} t(M1)(M2)(M3)(M4)(M5)$$

$$= \sum_{s,t} t(M2)(M5)(M1)(M3)(M4) \qquad \ldots (1.6)$$

where $(M2)(M5) = P(S_s|I_i).P(T_t|I_i.S_s.M_m)$

$= P(S_s|I_i).P(T_t|I_i.S_s)$, by Assumption 2

$= P(S_s.T_t|I_i)$, by Bayes' rule.

ASSUMPTION 3: $P(G_g|I_i.S_s) = P(G_g|I_i)$

$P(M_m|I_i.S_s.G_g) = P(M_m|I_i.G_g)$

Assumption 3 says merely that overnight stay destination is not affected by the number of annual overnight stays. This is clearly a simplification - one would expect, for example, that surfing officionados have a different pattern of site preference to the less enthusiastic. Nevertheless this assumption effects a considerable reduction in the complexity of (1.6):

$$E = (M1)(M3)(M4) \sum_{s,t} st P(S_s . T_t | I_i)$$

$$= (M1)(M3)(M4) E(T | I_i) \qquad \ldots (1.7)$$

Equation (1.7) is the form in which the model was estimated. $E(T|I_i)$ is a readily recognisable quantity - the expected activity, trips undertaken, while on holidays. Thus Equation (1.7) has a plausible interpretation: -

(i) (M1) is the proportion of individuals who participate in the activity,

(ii) $E(T|I_i)$ is the number of trips generated by each participant whilst on holiday,

(iii) (M3) is the proportion of holiday trips directed to region g, and

(iv) (M4) is the proportion of holiday trips directed to site m (see Equation 1.2).

Equation (1.7) is thus exactly equivalent to Equation 1.2 of this report and as presented in the Geelong Recreation Study - Phase 1 Report.

Simplifications in Model Specification

During the modelling a number of additional assumptions were made to simplify what was potentially a huge calibration task. This involved discarding variables and models which our increasing empathy with the data suggested were unlikely to be of value. In any case the starting point is indicated in Table 1.12. Note that SES stands for the vector of socio-economic descriptors.

Models M3 and M4 have one thing in common - the dependent variables are of the form $P(A|B)$ where B can take a large number of mutually exclusive values $B_1, B_2, \ldots$ For example, $B_1=1$ if the observation is of a surfer, $B_1= 0$ otherwise. In both cases our approach to this family of models was to first consider $P(A)$ as a function of both SES variables and indicator variables of $B_1, B_2, \ldots$ Thus if

$$P(A) = f(SES, I_1, I_2, \ldots)$$

then we may take $P(A|B_1) = f(SES, B_1, 0, 0, \ldots)$.

For example, in case M4, instead of estimating a separate (activity i) overnight stay site choice mode for each "activity, activity level" pair (granted that the region is g) we estimate a single overnight stay choice model which contains as dependent variables indicators of participation in activity i as well as the possibility of engaging in activity i at the site. Naturally this entails certain assumptions about the model parameters - for example that the marginal disutility of travel cost is the same for different holiday activity choices - nevertheless the potential reduction in the modelling task was too inviting to disregard.

We shall not discuss the specifications of these models but refer the reader to the previous section for a general description. The keener reader is referred to the Geelong Recreation Study - Phase 1 Report.

The Home-Based Model

The home-based model is considerably simpler in structure and a brief discussion will suffice to describe it. The events are set out in Table 1.13.

TABLE 1.12 - THE FORM OF THE MODELS

Model	Form	Independent Variables
M1	LOGIT	SES, locational factors
M3	LOGIT	Regional characteristics, SES, locational factors
M4	LOGIT	Site characteristics, SES, locational factors
$E(T \mid I_i)$	LOG-LINEAR	SES, locational factors

TABLE 1.13 - THE HOME-BASED TRIP EVENT SPACE

Event	Description
I_i	the individual participates in activity i
Q_q	the individual undertakes q annual home-based trips
R_r	the individual chooses region r for a home-based trip
W_w	the individual chooses site w for a home-based trip

TABLE 1.14 - THE HOME-BASED MODEL

Model	Dependent Variable
M1	$P(I_i)$
M8	$P(Q_q \mid I_i)$
M9	$P(R_r \mid I_i . Q_q . W_w)$
M10	$P(W_w \mid I_i . Q_q)$

TABLE 1.15 - MODEL SPECIFICATION

Model	Form	Independent Variables
M1	LOGIT	SES, locational factors
M9	LOGIT	Regional Characteristics, locational factors
M10	LOGIT	Site characteristics, locational factors
$E(V \mid I_i)$	LOG-LINEAR	SES, locational factors

The dependent variable is $E(V_w|I_i)$, the expected number of home-based trips to site w by a given participant in activity i. As before

$$E(V_w|I_i) = \sum_q (M1)(q)(M8)(M9)(M10) \qquad \ldots (1.8)$$

where the models are as set out in Table 1.14.

As before we assumed the following:

$$P(R_r|I_i.Q_q.W_w) = P(R_r|I_i.W_w),$$

$$P(W_w|I_i.Q_q.R_r) = P(W_w|I_i.R_r),$$

whence (1.8) becomes:

$$\begin{aligned} E(V_w|I_i) &= (M1)(M9)(M10)\Sigma qP(Q_q|I_i) \\ &= (M1)(M9)(M10)E(V|I_i) \end{aligned}$$

where V is the annual number of home-based activity trips made by each given individual (c.f. Equation 1.1).

The home-based model was estimated in this form with the independent variables as set out in Table 1.15. This is equivalent to equation 1.1 of this report and as presented in the Geelong Recreation Study - Phase 1 report.

For details of the final specification of these models the reader is referred to the previous section.

THE GEELONG RECREATION STUDY DATA

The data used in calibrating the Geelong models were collected from nearly 1800 individuals from 800 households in the urban areas of Melbourne, Geelong and Ballarat. Each respondent was required to recall all holidays and outdoor recreation activities undertaken over the previous 12 months.

The most obsious question to ask at this point is related to how accurately an individual can recall events undertaken over a 12 month period. There is some evidence to suggest that respondents tended to overestimate the number of times they went to particular sites if these sites were visited infrequently. For example, when shown a map of the Geelong Region which shows the You Yangs National Park there was a tendency for respondents to claim they had been to the You Yangs once in the past 12 months, when they may have been only once in the past 2 or 3 years. That is, there is subconscious tendency for occasional visits to a site to be brought forward to the past 12 months. It is difficult to see how such over-estimating can be avoided short of asking respondents to keep annual recreation diaries, or interviewers spending more time probing each reply to ensure a more accurate response. If the survey technique used in this study is repeated interviewers should be instructed to probe for accuracy particularly when respondents claim to have been to a site only once in the past 12 months.

It seems reasonable to assume that the extent of such response over-estimation decreases as the number of times an individual reports to have visited a site increases. If a respondent claims to visit the You Yangs every weekend, then one can safely assume that this is more or less the case.

Evidence from surveys undertaken by the Ministry for Tourism indicates that people tend to underestimate the number of holidays undertaken when asked to recall over a 12 month period. (A "safer" period is reported to be a one month period.) The extent of the underestimate has not been analysed.

Very little is known about how important these biases are in their effect on recreation demand analyses of the sort attempted in the Geelong Recreation Study. It may be that the overestimates associated with site visit recall is compensated by the underestimate of holiday participation. The Geelong Recreation Study's estimate of the number of visitors to the You Yangs indicates that the data is producing overestimates of visitor rates (possibly by as much as 50% in some instances).

A further problem associated with the sample of 800 households is related to lack of information on activities undertaken by only a small proportion of the population. Many such "minority" activities are important as they happen to be activities which can seriously affect the environmental stability of an area. If, for example, only 20 of a sample of 2000 individuals reported to have undertaken dune buggy activities, then very little rigorous work could be undertaken to "model" dune buggy behaviour. More information could be gathered by interviewing only dune buggy drivers on-site. This additional information can be used for modelling purposes provided the new observations are weighted by non dune buggy participants such that the observation set used for the modelling has the same proportion of dune buggy participants as that observed in the random sample (20 in 2000, or 1%)[1].

---

(1) For an example of the way this might be done, see S.R. Lerman, C.F. Manski and T.J. Atherton (1976) "Non Random Sampling in the Calibration of Disaggregate Choice Models", Department of Transport, Report No. DO-6-3-0021.

CHAPTER 2 - THE NEED FOR ATTITUDINAL DATA

INTRODUCTION

Because disaggregate models are calibrated on the observed patterns of recreation choice of individuals, they are commonly referred to as being behavioural. To the extent that recreation choice had been modelled against certain socio-economic and site characteristics it is possible to say that the study has produced models of recreation behaviour which are suitable for use in policy analysis.

However, the models are not behavioural if by that term we mean that individual attitudes and perceptions of choice attributes underlie the modelling process.

The Geelong Recreation Study approach relates individual characteristics and individual behaviour directly without reference to any underlying psychological motivations. Individuals have attitudes to each component of the recreation experience, and it is these attitudes that ultimately determine individual behaviour. Quite clearly, the more insight one has into the psychological factors motivating individual behaviour the more sensitive to policy management factors will be the final modelling process.

At the time of formulating the Geelong Recreation Study the subject of attitudinal analysis with respect to travel demand had not been thoroughly researched. The past 18 months has seen a growing body of theory and practice in the field of transport planning which demonstrates that attitudinal data can be successfully used to improve the reliability of transport planning models. In particular, Recker and Golob[1] used attitudinal

---

(1) W.W. Recker and T.F. Golob (1976) "An Attitudinal Model of Mode Choice", Transportation Research, Vol. 10, No. 5, October 1976 - an abstract of which is presented in the bibliography at the end of this chapter.

ratings by survey respondents of descriptive attributes of travel models in an attempt to improve the "traditional" model choice modelling techniques. Such attributes included features such as comfort, reliability of service, safety, opportunity to read, etc. Recker and Golob conclude that:

> Goodness of fit summaries indicate model performance better than that obtained using conventional logit models based upon perceived time and cost data.

We wish now to address the problem of how attitudinal data could be included in any future recreation demand study. It is suggested that future surveys should collect two types of information:

(a) Data suitable for reproducing the "conventional" analysis of the Geelong Recreation Study; viz, individual socio-economic and recreation travel information.

(b) Attitudinal data which will be of use in providing a description of the psychological "building blocks" which link individual socio-economic characteristics and reported recreation behaviour. (Conventional analyses bridge these "building blocks" with the one structure (the single regression of logit analysis), and in so doing provide no insight into the underlying decision making process of the individual.)

A THEORETICAL FRAMEWORK

It is important that the theoretical structure of the proposed motivational building blocks be understood at the outset to ensure that the right type of questions are included in any future survey. It is for this reason that we turn to a discussion of the structure of the recreation choice process.

The theoretical basis is Anderson's Information Integration theory(1) as proselytised, in particular, by Louviere(2). This approach will place any future analysis within a rigorous theory of perception and cognition and one wherein great emphasis is placed on developing true specifications of the underlying psychological laws. By this means it is hoped that the behavioural content of any data analysis will be enlarged both with the inclusion of a new socio-economic variable and a more sensitive analysis of the old.

Attitudes to Activities

The way in which the attitudinal data will be included in the analysis of recreation travel demand is illustrated in Fig. 2.1.

In Boxes 1 and 2 we are searching for a statistical relationship between individual socio-economic characteristics and attitudes to particular activities. To clarify the discussion we need to define two terms:

(i) An ACTIVITY is an outdoor recreation pastime, such as surfboard riding, picnicking, trailbike riding, etc.

(ii) An ATTRIBUTE of an activity represents an inherent quality of an outdoor recreation experience. For example, an outdoor recreation pursuit can be classified according to the attributes of (a) activeness or passivity, (b) being an activity undertaken with friends or whilst alone, (c) being dangerous or otherwise, etc.

---

(1) The keen student is referred to a series of Technical Reports by Anderson (1976), Center for Human Information Processing, University of California, La Jalla.
(2) A number of Louviere's studies is referred to in the bibliography at the end of this section. See also "Application of Psychological Measurement", J.J. Louviere, Center for Behavioural Studies, University of Wyoming, March 1977.

1. Individual (i) attitudes to attributes (X) of outdoor recreation activities are determined by socio-economic factors (SES)

   i.e. $\beta^i_X = \beta_X(SES_i)$

2. An individual's attitude to a particular activity (A) is determined by attitudes to the salient attributes of activities (of which there are T)

   i.e. $\alpha^i_A = \alpha_A(\beta^i_1,\dots,\beta^i_T)$

3.* The probability of an individual i participating in an activity A is jointly dependent on the socio-economic characteristics of i and his attitude to A:

   i.e. $P(\text{i participates in A}) = P(\alpha^i_A, SES_i)$

4.* Given that an individual will undertake activity A, the number of activity trips to site S is dependent on attitude to S *for activity* as well as person type (possibly only distance to S)

   i.e. $N^i_A(S) = N_A(\alpha^i_A(S), SES_i)$

5. Attitudes towards attribute (X) at site (S) determined by the physical supply of "attributes" at that site and, perhaps, person type

   i.e. $\delta^i_{A,X}(S) = \delta_{A,X}(P(S,X), SES_i)$

6. Attitudes to undertaking activity A at site S are determined by an individual's attitude to the site as supplier of facilities for the salient attributes

   i.e. $\alpha^i_A(S) = \alpha_A(\delta^i_{A,1}(S),\dots,\delta^i_{A,T}(S))$

* Boxes 3 and 4 are essentially the Geelong Recreation study model.

FIG. 2.1 - THE GEELONG RECREATION MODEL STRUCTURE MODIFIED TO INCLUDE ATTITUDINAL DATA

It is hypothesised that an individual's attitude to an activity is nothing more than his combined attitudes to the attributes which he perceives as being salient to the activity. That is, if an individual perceives surfboard riding as being (i) very active (ii) involving a lot of time with friends (iii) involving a certain amount of danger and, (iv) a healthy pastime, then those 4 attributes represent his salient set. It is the individual's attitudue to these 4 salient attributes that determine his attitude to surfboard riding.

The information required as a basis to analysing these attitudes could be gained by asking a set of questions similar to the following:

Question A: What is your attitude to the following outdoor recreation characteristics (Circle appropriate score.)

		Strongly Dislike			Indifferent			Strongly Like
1.	ACTIVE PASTIMES	1	2	3	4	5	6	7
2.	PASSIVE PASTIMES	1	2	3	4	5	6	7
3.	BEING WITH FRIENDS	1	2	3	4	5	6	7
4.	BEING BY YOURSELF	1	2	3	4	5	6	7
5.	DANGEROUS PASTIMES etc.	1	2	3	4	5	6	7

A complete list of possible attributes could be determined in a pilot survey.

The information gained from this question will provide the basis for developing insight into the analysis described in Box 1 of Fig. 2.1. In fact this question measures $\beta_X{}^{L}1$.

Question B: How do you rate your attitude to the following outdoor recreation activities?

		Strongly Dislike			Indifferent			Strongly Like
1.	SURF SWIMMING	1	2	3	4	5	6	7
2.	PROTECTED BEACH SWIMMING	1	2	3	4	5	6	7
3.	SURFBOARD RIDING etc	1	2	3	4	5	6	7

This information will enable the choice processes illustrated in Box 2 of Fig. 2.1 to be explored. That is, combined with the information from Question A, the data can be analysed to provide an insight into individual attitudes to a particular activity as related to his attitudes to the activity's salient attributes. In fact, this question directly measures α_A^i.

It is useful to collect information on how individuals perceive activities with respect to the salient attributes of the activity. To do so the following set of questions would have to be asked:

<u>Question C</u>: For Activity 1 (surf swimming for example). Presented below is a list of possible characteristics associated with surf swimming. We are interested in how you rate surf swimming with respect to each of these attributes:

Surf Swimming is (a) not an active pastime

highly active

1 2 3 4 5 6 7

(b) not a passive pastime

highly passive

1 2 3 4 5 6 7

(c) not a pastime associated with danger

a highly dangerous activity

1 2 3 4 5 6 7

etc.

This question could be repeated for each activity under consideration. It has no place in the model structure outlined in Figure 2.1.

Attitudes to Activity Sites

We turn now to Boxes 5 and 6 in Fig. 2.1 and discuss the type of data required to interpret individual attitudes to sites. Box 5 is concerned with individual attitudes to a site for undertaking particular activities. That is, for each activity the following type of question is asked of respondents who have undertaken the activity at least once.

Question D: How do you rank each of the following sites as places to undertake surfswimming (or whatever)

		worst possible place to surf					perfect place to surf		Don't know
1.	PORTSEA		2	3	4	5	6	7	8
2.	PT LEO	1	2	3	4	5	6	7	8
3.	etc.	1	2	3	4	5	6	7	8

This question measures $\alpha^i_A(S)$.

Examination of Box 6 indicates that information regarding individual attitudes to particular activity attributes at each site would be required to provide complete information on site attitudes. It is proposed that this data be collected in the following way:

Question E: How do you rank the following characteristics of PORTSEA as a place to undertake Surf Swimming? (This question is updated for each site and each activity undertaken.)

		most unreliable surf						always good surf
1.	RELIABILITY	1	2	3	4	5	6	7
		always filthy						always clean
2.	CLEAN BEACH	1	2	3	4	5	6	7
	etc.							

This question measures $\delta^{i}_{A,X}(S)$.

THE ATTITUDINAL DATA IN THE CONTEXT OF THE TOTAL ANALYSIS

In the foregoing discussion we have examined a theoretical basis for attitudinal analysis and described the type of questions which would have to be asked to obtain a suitable data base for such an analysis. From an analysis of the attitudinal data we will have an understanding of two phenomena:

(i) Expected levels of individual attitudes to activities with respect to the socio-economic characteristics of the individual - α^{i}_{A}. An attitude to a single activity (A) will be expressed as a combination of attitudes to the salient attributes of that activity.

(ii) An individual's attitude to a particular site (S) for undertaking activity A - $\alpha^{i}_{A}(S)$. These attitudes could be related to the physical attributes of the site.

It is important to appreciate that the main body of any future survey would still be concerned with collecting socio-economic and recreation participation data of the sort collected and analysed in the Geelong Recreation study. Indeed, Boxes 3 and 4 in Fig. 2.1 are nothing more than the Geelong Recreation Model with α^{i}_{A} and $\alpha^{i}_{A}(S)$ as new sets of variables.

ANNOTATED BIBLIOGRAPHY

The following references have been chosen to represent a cross-section of recent work in attitudinal modelling. Most of the work reviewed is concerned with transport mode-choice modelling. Whilst many recreational[1] surveys have included attitudinal type questions none has done so in a way which has produced information of a sort suitable for inclusion in any theoretical model of behaviour.

(1) See for example McKenry (1975) Recreation, Wilderness, and the Public, Victorian Department of Youth Sport and Recreation, which includes some useful attitude responses.

AUTHOR Ricardo Dobson

TITLE: "Towards the Analysis of Attitudinal and Behavioural Responses to Transportation System Characteristics".

SOURCE: Transportation (4) 1975.

ABSTRACT: The paper is concerned with the status of contemporary attitudinal data collection. The first part of the paper gives a valuable classification of the attributes of transportation systems and the various transportation system options which planners might consider. This classification could be of value in determining the structure of attitudinal surveys.

The second and more valuable section of the paper gives a classification of survey questions into a number of tasks:

1. The ranking task. The respondent ranks modes with respect to various attributes.

2. The paired comparison task. The respondent ranks pairs of desirable attributes.

3. The pick k of n-1 task. Given a fixed attribute, the respondent classifies a list of attributes into "similar" and "not similar".

4. Semantic differential. With respect to a given attribute, e.g. comfort, two poles are identified (e.g. comfort, discomfort) and the respondent ranks the mode between these extremes according to a given scale.

5. Category scaling task. Attributes are ranked on a scale ranging from "extremely important" to "extremely unimportant".

Dobson claims that the latter three provide more robust results than the former two. In particular, he strongly recommends that similarity data (category 3) be collected.

The paper is written in a style that manages - miraculously - to be simultaneously both gaudy and turbid. Nevertheless, the work achieves a notable classification of attitudinal survey techniques and should prove of considerable value to those involved in survey design.

AUTHORS: Hensher, D.A., Macleod, P.B. & Stanley, J.K.

TITLE: "Usefulness of Attitudinal Measures in Investigating the Choice Travel Mode"

SOURCE: International Journal of Transport Economics, Vol. II, April 1975.

ABSTRACT: The paper is concerned with developing measures of travellers' attitudes under four headings conveniently summarised as:

Time
Cost
Comfort
Convenience

The data exercise was gathered in the suburbs of Sydney, Australia in 1973. Two types of questions were asked:

1. Importance Points. Here respondents assign points from 0-100 to a given attribute.

2. Constrained Importance. In this case respondents assign a total of 100 points to a number of attributes.

The second method has the advantage that respondents are forced to tradeoff attributes, and a disadvantage that the number of attributes that can be considered simultaneously in this manner is limited to about 5. These measures are not quite as substitutable as might be hoped: the correlation ranges from .82 to .89.

The paper attempts to develop models of the overall response to comfort and convenience as dependent on the component attributes of these descriptors. For example, the response to comfort is modelled as a linear combination of responses to seating quality, seating quantity, the possibility of relaxing, cleanliness, ease of travel, being out of the weather, no congestion, and ventilation. The comfort model appears to be much more satisfactory than the convenience model. The comfort model has all variables with the right sign and the worst t value is 1.0. On the other hand the convenience model is very bad: no variables are significant at the 95% level and the overall R^2 is very low. Louviere (op.cit.) claims that overall response should be modelled as a product of attribute responses - this may be the source of the problem.

The measure of response to a mode is a sum of the perceived levels of satisfaction associated with each attribute weighted by the importance of that attribute. The authors also consider another method of aggregating the coded response, equivalent to setting the perceived level of satisfaction to zero if the rival mode dominates the main question in this regard. No discussion of this curious procedure is given but it appears to reduce multi-collinearity in certain models and to increase R^2 in others.

In the final section of the paper the authors used the attitudinal measures in a logit-choice model. Whereas the calibrated model is no more accurate a predictor of mode choice than models calibrated with the usual physical variables of cost and time, the authors argue (with justification) that they have managed to include in the model the variables which are most significant for policy applications.

The paper is very good and has not received the attention it deserves. Perhaps this is due to its emphasis on cardinal utility which places the work in the context of microeconomics, whereas the more natural place would seem to be behavioural psychology. The latter half of the work is discussed from the point of view of the "signed preference indicator" which has been discussed above. The authors do not make out a telling case for this concept and it is hard to resist the idea that it has distracted them from asking a number of interesting and important questions about traveller perception.

AUTHOR: Levin, R.P.

TITLE: "The Development of Attitudinal Modelling Approaches in Transportation Research"

SOURCE: The Third International Conference on Behavioural Travel Modelling, 1977, Tanunda, South Australia.

ABSTRACT: The paper begins with a review of attitudinal modelling and in this regard contains a pertinent bibliography. In particular, Levin is optimistic about the uses of Anderson's information integration theory. A laboratory experiment in the manner of Louviere (loc.cit.) is discussed: the subject is driver perception of safety as a function of weather, time of day, speed, and hours of consecutive driving. The paper does not go into the analysis in great detail, but is of value in that it describes a different sort of application to those described by Louviere.

AUTHOR: Louviere, J.J. et.al.

TITLE: "Theory of Empirical Evidence in Real-World Studies of Human Judgement".

SOURCE: Centre for Behavioural Studies, University of Wyoming, Laramie.

ABSTRACT: This paper studies the relationships between the various stages of perception and action in the context of consumer response and judgement. A number of models are identified:

1. Physical stimulus. This model has the form

$$x = f(X)$$

where X is some physically measurable quantity (stimulus) and x its subjective counterpart.

2. Evaluation. A number of subjective qualities are combined in a single evaluation. The form may be

$$I = \Sigma x \text{ or } I = \Pi x$$

where x represents a number of the measures from I and I is an unobserved dimension of subjective evaluation - something like utility in economic theory.

3. The overall response. This model identifies the relationship between the psychological response and the subjective unobserved evaluation I:

$$R = R(I)$$

4. The observed behaviour. This model translates the overall response into observed behaviour

$$B = B(R).$$

For models 1, 3, and 4 the functional form

$$Y = a+Bx^{s}$$

is employed. The evidence suggests that the product form in 2 is correct. In a typical study (3 are presented) the variables might be:

X range of products in a store

x rating on a 0-100 scale of the range of a store

I sum or product of x terms

R overall assessment on a 0-100 scale

B sales at store

The work is very interesting and the results very good given the data. Studies of this sort could be mounted without a great deal of effort.

AUTHOR: Louviere, J.J. et.al.

TITLE: "Attitudes, Attitudinal Measurement and the Relationship Between Behaviour and Attitudes".

SOURCE: Third International Conference on Behavioural Travel Modelling, 1977, Tanunda, South Australia.

ABSTRACT: The paper discusses two case studies of Psychological Modelling of Choice Behaviour from the standpoint of Anderson's algebraic integration theory of perception. The case studies are:

- choice of town as residence by Industrial Employees
- choice of shopping destination.

In both cases the primary data source (from which the model is calibrated) derives from controlled laboratory experiments, typically ratings and selections of alternatives by University students. Correspondence with observed behaviour is much stronger than non-psychological models.

The paper is very persuasive about the applications of the methodology to travel modelling and the authors are probably onto something. It is worth emphasizing how extremely cheap the data is to collect: the analysis/information-gathering ratio is very high. The work is also valuable in that it gives an exhaustive account of the "tricks of the trade", that is, the way the models are specified, estimated, analysed, respecified and so on.

AUTHOR: Louviere, J.J. et.al.

TITLE: "Applications of Information Processing Theory to the Analysis of Urban Travel Demand"

SOURCE: Environment and Behaviour, Vol. 9, March 1977.

ABSTRACT: This paper reports some experiments by the authors in assessing consumer perception of public transport from the point of view of Anderson's Information Integration Theory.

The experiments are laboratory based, that is, models of perception are calibrated on data obtained by asking respondents (usually students) to respond to various factorial designs of levels of service.

In the first experiment, respondents were asked to rate on a 0-100 scale the 27 possibilities arising from three levels of fare, three levels of frequency, and three proximities to bus stops. A multiplicative form was assumed for the overall response to the mode and the authors claim an excellent fit was achieved. One result that casts a doubt concerning the methodology is that fare was found to have the greatest influence - this would seem to run counter to almost universal experience in revealed preference. The authors also report than an attempt to include a greater number of bus attributes in the model was unsatisfactory.

The authors also report a similar experiment where overall response is modelled as dependent on fare, reliability, frequency, and a cross-elasticity term in the price of gasoline. A slightly more complex

form of the multiplicative model was used - in this case factors could appear raised to powers. On the face of it the models again overemphasise fare compared to other levels of service.

The final experiments describe models similar to the above calibrated across trip purposes: work and shopping. Once again, there are some disturbing non-correspondences with results obtained by revealed preference methods.

This is an important paper because obtaining information on traveller perception of transportation systems by laboratory experiments has enormous significance for transportation modelling if it can reproduce observed system responses. The methods described in this paper seem extremely promising. Nevertheless the correspondence with revealed preference seems, on the face of it, to be less than perfect. There is great scope for further research.

AUTHOR: Recker, W.W. and Golob, T.F.

TITLE: "An Attitudinal Modal Choice Model"

SOURCE: Transportation Research, Volume 10, No. 5, October 1976.

ABSTRACT: The paper develops an attitudinal modal choice model for work trips. There are three steps in the modelling procedure:

1. Individuals are first clustered into groups facing a reasonably homogeneous choice environment. The clustering technique employed was the ISODATA algorithm developed by Ball and Hall. Five market segments were identified. However, the authors only had data to estimate robustly two of these.

2. Factor analysis is used to identify the latent dimensions of commuter choice. The important factors appear to relate to:

   . service
   . vehicle ride quality
   . crowding
   . convenience
   . performance
   . personal environment
   . sociability.

   Principle component analysis with VARIMAX rotation was used.

3. A logit utility model is developed with the factors identified in the second part of the work used as variables.

   It is hard to resist the notion that attitudinal modelling is a major new direction in transportation analysis. It seems likely that this paper will have a seminal role in such a development.

AUTHOR: Recker, W.W. and Golob, T.F.

TITLE: "A Behavioural Travel Demand Model Incorporating Choice Constraint".

SOURCE: The Research Laboratories of General Motors Corporation, Warren, Michigan.

ABSTRACT: This paper is a first version of the paper 'An Attitudinal Modal Choice Model' that the authors subsequently published in the October, 1976 version of Transportation Research.

AUTHOR: Recker, W.W. and Golob, T.F.

TITLE: "An Attitudinal Mode Choice Model".

SOURCE: Research Laboratories, General Motors Corporation, Warren, Michigan.

ABSTRACT: This paper is an early version of the paper of the same name appearing in Transportation Research (October, 1976).

AUTHOR: Golob, T.F. and Recker, W.W.

TITLE: "Attitude Behaviour Models for Public Systems Planning and Design".

SOURCE: Research Laboratories, General Motors Corporation, Warren, Michigan.

ABSTRACT: The content of this work is a preliminary version of the work published in the author's Transportation Research paper: An Attitudinal Modal Choice Model (October, 1976).

AUTHOR: Stopher, P.R.

TITLE: "On the Application of Psychological Measurement Techniques to Travel Demand Estimation"

SOURCE: Environment and Behaviour, Vol. 9, March, 1977.

ABSTRACT: This is a review article aiming to bring the possibility of psychometric investigation of traveller behaviour to the attention of psychologists. Stopher makes two points of interest:

1. If psychometric analysis of attitudinal data is to be of importance to transport planners, then the physical correlates of attitudes will need to be analysed. Presumably this is a call to develop models of subjective response determined by objective stimuli.

Note Louviere and his associates appear to be doing this. (See for example "Theory of Empirical Evidence" (loc.cit.).)

2. Stopher also warns that those involved in this area of transportation research may be unqualified to make refinements to psychometric techniques required.

AUTHOR: Thomas, K.

TITLE: "A Reinterpretation of the Attitude Approach to Transport Mode Choice and an Exploratory Empirical Test".

SOURCE: Environment and Planning 1976, Volume 8.

ABSTRACT: The paper propounds a theory of cognitive structure from which viewpoint the transport mode choice is studied. The theory distinguishes two determinants of overall attitude towards objects or behaviour. These are:

1. The strength of the belief concerning the attributes of that object or behaviour. For example, if the behaviour considered is "using the bus" then the attribute might be "carrying heavy shopping" and the respondent might attach a certain likelihood to that outcome given the behaviour.

2. An evaluation of the attribute of the object or behaviour. In the above example the respondent would be asked to attach a value weight to carrying heavy shopping.

The model underlying the theory holds that the overall attitude towards the object or behaviour is now the product of 1 and 2, perhaps summed over all salient attributes of the object or behaviour.

Thomas warns that failing to distinguish beliefs and attitudes at the survey design stage can lead to information of doubtful value being collected. An example cited by Dobson (op.cit.) is criticised for this practice.

Thomas' theory is tested on a group of women shoppers in the south east of England. The results appear to have been statistically significant, but rather less satisfactory than similar work in other areas. The author suggests that the reason may be that attitude to transport may be more influenced by behavioural commitment than in other areas of consumer demand.

An interesting section of the work concerns the stability of the set of attitudes over a period of change in the level of service provided by public transport. In fact, on a lot of occasions changes were slight, but a significant effect on beliefs was observed. On the other hand, no significant change in overall attitudes was measured.

Thomas is quite persuasive concerning his theory of cognitive structure. Moreover, one cannot help feeling that only rudimentary analytical and statistical techniques have been brought to bear - the scope for innovation is great. In any case, the theory as it has so far developed is capable of wide applications.

CHAPTER 3 - FACTORS IN RECREATIONAL BEHAVIOUR

DIMENSIONS OF BEHAVIOUR

The Geelong Recreation survey did not collect attitudinal data. Nevertheless objective data - socio-economic and trip making behaviour - can be used to cast considerable light on recurrent patterns of trip making and, implicitly, on the attitudes underlying those patterns. In this chapter Factor Analysis is used to determine seven dimensions of behaviour which explain nearly half of the variance in our data set. These factors are readily identifiable with behavioural or socio-economic archetypes and seem quite robust with respect to factor methodology.

FACTOR ANALYSIS

To enable this paper to be read by those with minimal statistics we shall begin with a brief and informal discussion of factor analysis. Broadly speaking given a survey which involves a number of random variables x_1, x_2 ... (eg. Age, Income ...), a factor analysis seeks to uncover a limited number of usually uncorrelated variables f_1, f_2, ... of which each x_i can be expressed as a sum (to a good approximation). These variables f_i are called factors. Factor analysis has point only if:

- there are many fewer f_i than x_i, and
- the approximation to each x_i is good.

If the first fails then we may as well have stayed with the x_i; if the second fails then the f_i are an insufficient explanation of the x_i. The measure of degree of explanation of each x_i is its communality - the proportion of variance explained by the factor set.

The f_i must be constructed from the data set - in fact they are chosen to be sums of the x_i. This seems paradoxical - the x_i are sums of the f_i which are, in turn, sums of the x_i. In fact, the analysis is of limited value if one does not accept that factors are associated with real psychological components from which personality or behaviour is formed. For example, we shall deal with a factor which is positively associated with watching football and negatively associated with visiting historical centres. We shall argue that this factor is associated with a non participatory philistine dimension in the behavioural makeup of our sample - in fact we call it FOOTBALL-GOTHIC. A large part of factor analysis is just this - examining the manner in which each f_i is composed of the x_i and trying to fit some psychological or behavioural interpretation to it.

To dispel a misconception we emphasis that factor analysis does not partition a group of observations; rather, it isolates a number of descriptors, preferably few in number and certainly co-existent (to a greater or lesser extent) in all individuals. Thus it is not true that a factor describes an individual - a factor is merely part of a description involving many elements. It would be more correct to say that a factor described a behaviour or propensity to behave in a certain way, although even here care must be taken as different factors may produce similar behaviour (for example, a factor distinguishing in the ensuing analysis is associated with low-energy tourism and spectating: as with FOOTBALL-GOTHIC this factor is highly correlated with watching football). Moreover a given individual may be described by factors which pull in different directions. For example, we shall isolate two factors, one associated with water sports (YOUNG-SURF) and one associated with the reproductive stage of the life cycle (YOUNG-MARRIEDS). Membership of the 20-35 age group increases the weight of both these factors in determining behaviour - yet the former is associated with a high level of outdoor recreational activities and the latter with a low level. The point is simply that, all things being equal, membership of this age group increases the probability that one will participate in water sports, thereby increasing certain activities, as well as increasing the probability that one will be married with dependent infants, thereby decreasing these activities. If one wishes to observe the net effect one may always examine the direct correlation - the factor approach decomposes this correlation into two separate components.

An example of a factor analysis is set out in Table 3.4. The columns correspond to the factors, the rows to the original variables. In fact, the rows give the decomposition of each variable into a sum of factors. For example:

AGEI = .00939 FACTOR 1 - .00728 FACTOR 2 + ...

The entry corresponding to a given variable in a factor is called the loading of the variable in the factor.

Levels of Significance

Significance of Factors: There are a number of methods of judging the statistical significance of a factor, usually involving the eigenvalue - which may be interpreted as the total variance (ie. in all variables) accounted for by the factor. The simplest is Kaiser's rule - a factor is admitted if its eigenvalue is greater than unity. Figure 3.1 gives a plot of the eigenvalues of a factor analysis on 40 variables. It will be observed that by Kaiser's rule the first 16 factors will be accepted. A different criterion is obtained by arguing that, after a certain number, each additional factor accounts for only a random share of the remaining variance. This means that the eigenvalue plot becomes effectively linear. Some analysts have argued that only factors lying above this line should be accepted - usually with large data sets this leads to the extraction of a more conservative number of factors.

Significance of Loadings: It is of great interest, especially when seeking the behavioural dimension underlying a factor, to know whether a given loading is significant. As distinct from

FIG. 3.1 - EIGENVALUE OF POPULATION BEHAVIOUR AND S.E.S.

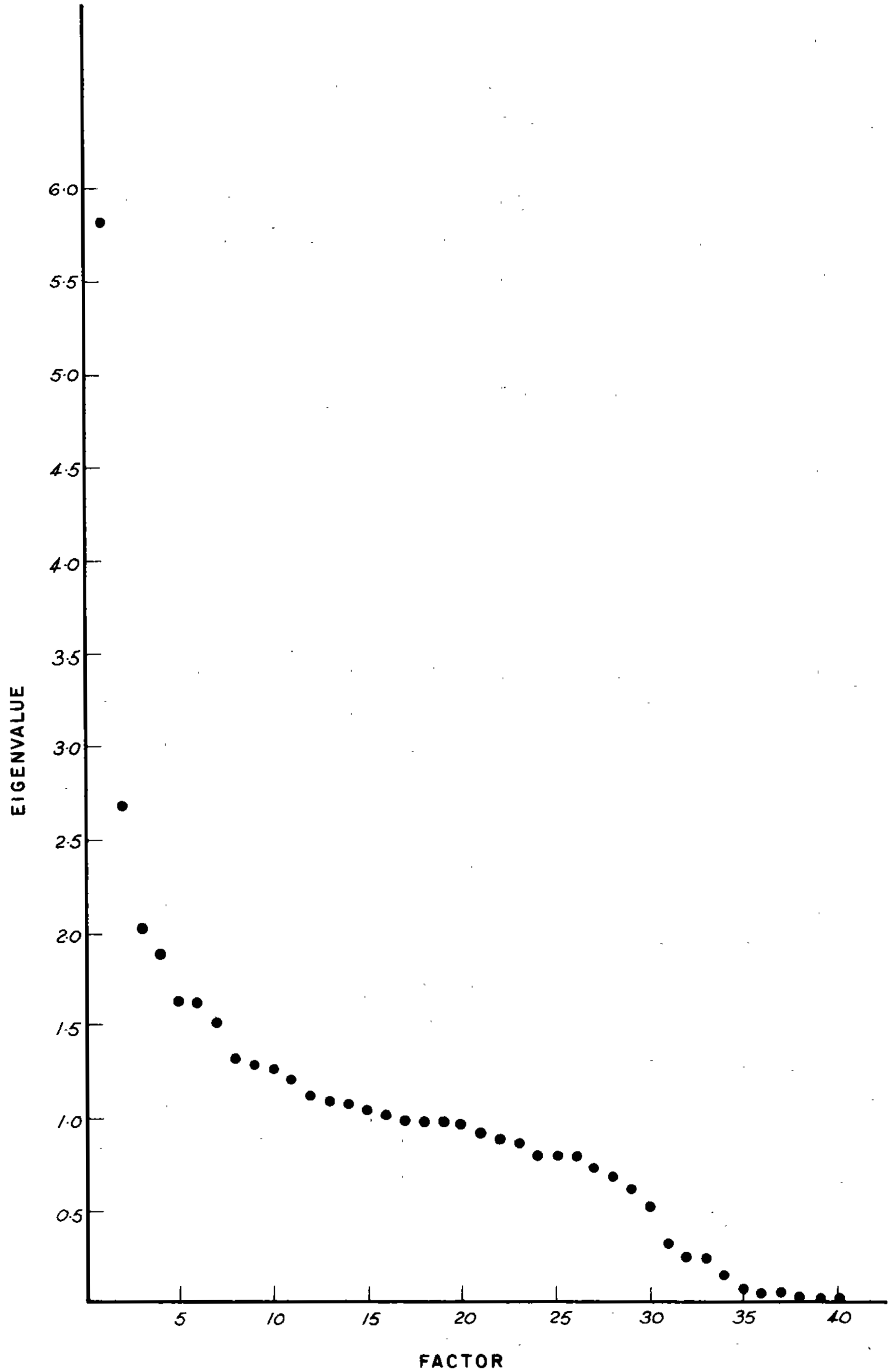
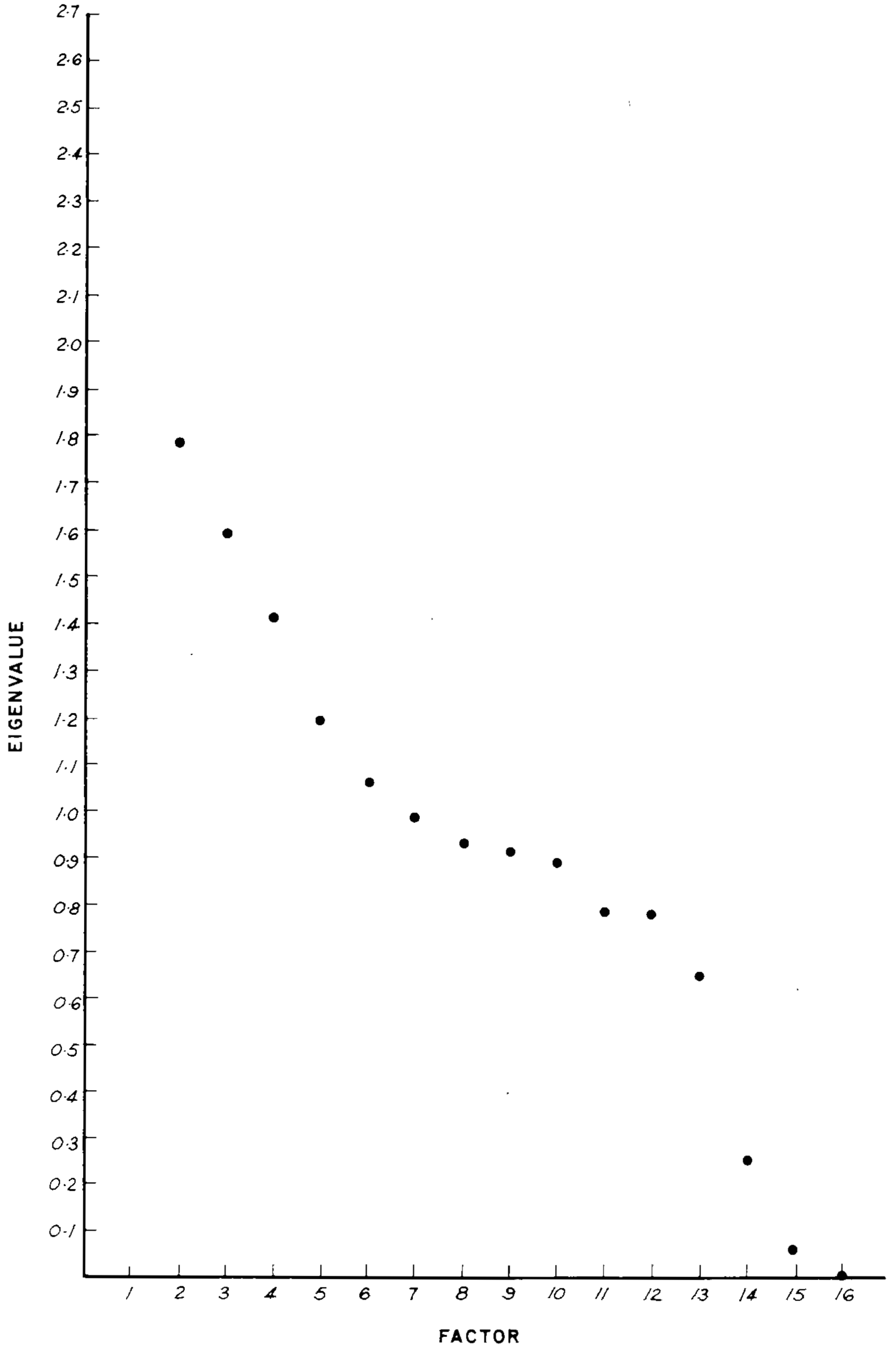


FIG. 3.2 - EIGENVALUE OF POPULATION S.E.S.



the factors themselves, loadings can be tested quite rigorously. Figure 3.3 presents a plot of the significance level for each of 40 factors at the 5% and 1% level. The loading of any variable may be considered significantly different from zero at the appropriate level if its magnitude exceeds the value given. Note that this variable depends on the factor number.

The Variables

The variables used in the analysis are described in Tables 3.1 and 3.2. The unit of observation is the individual; Table 3.1 gives the socio-economic descriptors and Table 3.2 enumerates the categories of trip making. Each observation consists of the vector of socio-economic data and a vector of trip frequencies. The correlation maxtrix of this variable set is given in Table 3.7.

The Factors: Statistical Significance

An OBLIMAX factor analysis was performed on the variables described in the preceeding section. A list of factors and eigenvalues is given in Table 3.3, and a plot in Figure 3.1. It will be seen that by Kaiser's Rule 16 factors are identified. On the other hand inspection of the plots indicates that there are at most 11 factors clearly distinguished from random explanation of variance. Of these, the last four are borderline and would seem to stand or fall together. The first seven factors are, however, quite clearly distinguished. Together, these seven factors explain 42.9% of the variance in the data.

The Factor Analysis

In this section we shall give a detailed account of each of the 11 factors. (See Table 3.5.)

Factor 1: The most significant behavioural loadings in factor are:

FIG. 3.3 - SIGNIFICANCE OF LOADINGS

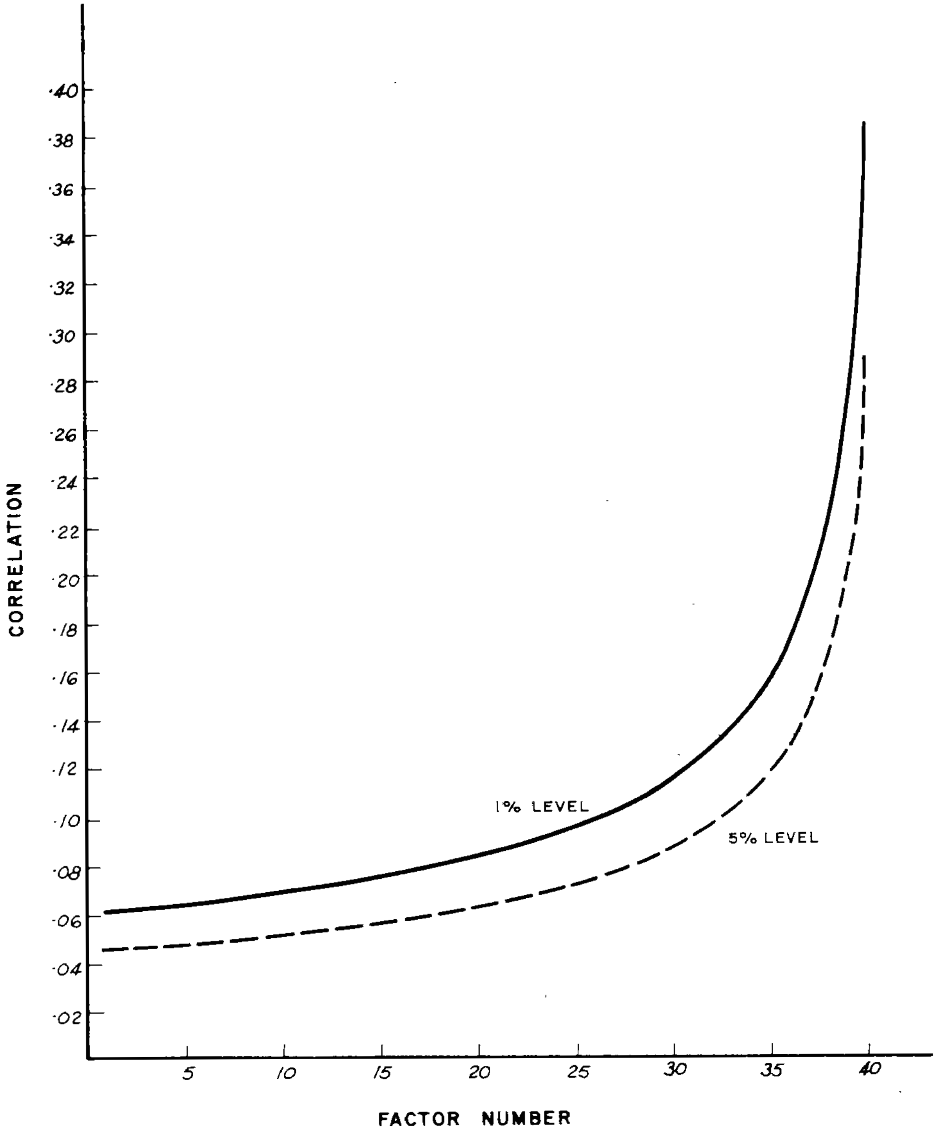


TABLE 3.1 - SOCIO-ECONOMIC VARIABLES

SEX	=	1 if individual is female
		0 if male
CSE	=	1 if individual has completed secondary education
	=	0 otherwise
OCC1	=	1 if individual involved in passive occupation
	=	0 otherwise
DBEACH	=	Distance (Km) to nearest protected beach from individual's place of residence
DSURF	=	Distance to nearest surf beach from individual's place of residence
DPARK	=	Distance to nearest "national park" from place of residence
AGE1	=	1 if individual aged 13-20 years
	=	0 otherwise
AGE2	=	1 if individual aged 21-34 years
	=	0 otherwise
AGE3	=	1 if individual aged 35-60 years
	=	0 otherwise
AGE4	=	1 if individual aged over 60 years
	=	0 otherwise
ADJINC	=	Adjusted income of individual
ANZ	=	1 if born in Australia or New Zealand
	=	0 otherwise
NCHLD	=	Number of children (5-13 years) in individual's income unit
NBABY	=	Number of babies (0-4 years) in individual's income unit
NA	=	1 if individual has no access to car
	=	0 otherwise
EMP	=	1 if individual unemployed
	=	0 otherwise

TABLE 3.2 - TRIP CATEGORIES: AN INDIVIDUAL'S "AGE" AND THE PATTERN OF PARTICIPATION IN OUTDOOR RECREATION ACTIVITIES

(Numbers from Survey)

	13-14 years	15-19 years	20-24 years	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	50-54 years	55-59 years	60-64 years	65-69 years	70-74 years	> 75 years
1. Visiting beach	45 (5.3)	103 (12.3)	81 (9.4)	127 (14.9)	76 (8.9)	98 (12.0)	40 (4.1)	54 (6.3)	65 (7.7)	35 (4.1)	38 (4.5)	28 (3.9)	17 (2.0)	16 (1.7)
2. Swimming - surf beach	32 (5.3)	84 (16.3)	86 (16.4)	92 (17.9)	58 (11.2)	44 (8.4)	38 (7.3)	17 (3.4)	28 (5.4)	15 (3.0)	15 (2.9)	4 (0.8)	1 (6.2)	3 (0.6)
3. Swimming - protected beach	33 (5.7)	87 (14.9)	88 (15.0)	91 (15.7)	58 (10.0)	48 (8.7)	43 (7.4)	33 (5.6)	29 (5.0)	10 (1.7)	14 (2.5)	8 (1.3)	4 (0.7)	7 (1.3)
4. Swimming - fresh water	38 (12.4)	57 (18.4)	51 (16.8)	42 (13.8)	21 (7.0)	35 (11.5)	21 (6.9)	7 (2.4)	12 (4.0)	6 (2.0)	6 (2.0)	1 (0.3)		
5. Rowing and canoeing	8 (13.6)	13 (22.1)	7 (11.9)	7 (11.9)	7 (11.9)	4 (6.8)	7 (11.9)	2 (3.4)	3 (5.1)			1 (1.7)		
6. Surfboard riding	14 (12.9)	33 (30.2)	19 (17.4)	21 (19.4)	3 (2.7)	4 (3.6)	5 (4.5)	2 (1.8)	4 (3.6)	1 (0.9)				
7. Water skiing - power boating	6 (3.7)	32 (19.9)	33 (19.9)	43 (26.6)	22 (13.1)	12 (7.4)	6 (3.6)	1 (0.6)	5 (3.0)			1 (0.6)		
8. Boat cruising	5 (3.2)	25 (15.7)	21 (13.1)	24 (15.7)	4 (5.8)	10 (6.3)	14 (7.7)	17 (10.7)	16 (10.0)	5 (3.1)	7 (4.5)	2 (1.2)	1 (0.6)	
9. Sailing	5 (9.1)	10 (18.2)	6 (10.8)	4 (7.2)	11 (19.9)	2 (3.6)	2 (3.6)	4 (7.3)	4 (7.2)	1 (1.8)	1 (1.8)	3 (5.5)		
10. Boat fishing and surf fishing	14 (6.2)	32 (14.3)	24 (10.7)	47 (20.8)	25 (11.1)	17 (8.8)	14 (6.2)	15 (6.6)	14 (6.2)	8 (3.0)	4 (1.7)	4 (1.7)	1 (0.4)	3 (1.2)
11. Bank/pier fishing	25 (7.0)	55 (15.6)	34 (9.5)	34 (9.5)	31 (9.3)	35 (8.8)	39 (10.4)	23 (6.3)	32 (9.0)	11 (3.1)	12 (3.3)	9 (2.6)	2 (0.6)	4 (1.2)
12. Skin diving, snorkel/scuba diving, spear fishing	5 (10.2)	8 (16.3)	8 (16.2)	14 (28.7)	6 (12.2)		3 (6.0)	1 (2.0)	3 (6.0)	1 (2.0)				
13. Bird watching, bush walking	42 (6.4)	98 (14.8)	81 (11.9)	101 (15.5)	48 (7.3)	50 (7.6)	52 (7.8)	40 (6.1)	43 (6.4)	20 (3.1)	34 (5.3)	18 (2.9)	5 (0.8)	8 (1.1)
14. Hiking, mountaineering	3 (4.7)	15 (23.4)	14 (21.9)	13 (20.4)	2 (3.2)	2 (3.2)	4 (6.3)	3 (4.8)	3 (4.8)	3 (4.8)		1 (1.6)		
15. Hunting	2 (2.0)	26 (25.6)	22 (21.8)	22 (21.8)	5 (5.0)	4 (4.0)	5 (5.0)		6 (6.0)	2 (2.0)		1 (1.0)		1 (1.0)
16. Motor cycle activities, dune buggies and other 4 wheel cross-country vehs.	4 (3.9)	21 (20.5)	19 (18.4)	23 (22.4)	11 (10.6)	7 (6.8)	4 (3.9)		5 (4.9)		2 (2.0)	1 (1.0)		
17. Site seeing, recreation driving	59 (4.4)	158 (11.7)	152 (11.3)	178 (13.3)	126 (9.4)	108 (8.0)	88 (6.6)	91 (6.8)	104 (7.8)	71 (5.3)	64 (4.8)	46 (3.4)	29 (2.2)	25 (1.7)
18. Picnicking and barbequing	55 (5.9)	120 (12.4)	112 (11.8)	148 (15.8)	98 (10.6)	85 (8.8)	84 (8.8)	57 (6.1)	68 (7.3)	30 (3.1)	31 (3.2)	19 (1.9)	8 (0.8)	10 (1.0)
19. Site seeing, recreation driving and barbequing picnicking	9 (2.9)	39 (12.6)	42 (13.6)	46 (14.9)	30 (9.7)	34 (11.0)	23 (6.4)	24 (6.5)	8 (2.6)	10 (3.2)	14 (4.5)	8 (2.4)	3 (0.9)	7 (2.2)
20. Golf	3 (2.0)	9 (6.7)	13 (8.5)	30 (19.8)	17 (11.2)	13 (8.5)	12 (8.0)	12 (8.0)	13 (8.5)	9 (6.7)	6 (3.9)	4 (2.6)	2 (1.4)	1 (0.7)
21. Visiting historical centres	38 (6.0)	66 (10.5)	54 (11.0)	102 (16.2)	56 (8.5)	55 (8.3)	51 (8.1)	45 (7.1)	43 (7.0)	25 (4.0)	30 (4.7)	17 (2.7)	8 (1.2)	12 (2.1)
22. Outdoor sporting activities	46 (10.1)	89 (19.6)	64 (14.0)	74 (16.4)	41 (8.9)	35 (7.7)	25 (5.4)	11 (2.3)	17 (4.0)	15 (3.4)	18 (3.3)	7 (2.3)	3 (0.6)	1 (0.2)
23. Outdoor sporting spectating	22 (5.3)	50 (12.6)	52 (13.2)	65 (16.3)	34 (8.7)	30 (7.6)	31 (7.9)	19 (4.8)	28 (7.1)	28 (6.0)	11 (3.7)	10 (2.9)	6 (1.9)	
24. Football spectating	16 (4.8)	51 (15.3)	47 (14.7)	55 (16.5)	29 (8.7)	23 (6.9)	20 (6.0)	21 (6.3)	22 (6.6)	20 (6.0)	11 (3.3)	10 (2.7)	5 (1.5)	2 (0.6)
Sample	82 (4.6)	212 (11.8)	190 (10.6)	230 (12.8)	152 (8.6)	141 (7.9)	112 (6.9)	229 (6.5)	145 (8.1)	92 (5.2)	92 (5.2)	74 (4.2)	24 (2.8)	21 (2.4)

NOTE Numbers in brackets represent the proportion of all undertaking the particular activity.

TABLE 3.3 - EIGENVALUES

Factor	Eigenvalue	PCT of VAR	CUM PCT
1	5.80796	14.5	14.5
2	2.68640	6.7	21.2
3	2.02474	5.1	26.3
4	1.88078	4.7	31.0
5	1.62369	4.1	35.1
6	1.60292	4.0	39.1
7	1.51933	3.8	42.9
8	1.31437	3.3	46.2
9	1.28009	3.2	49.4
10	1.25376	3.1	52.5
11	1.16840	2.9	55.4
12	1.11831	2.8	58.2
13	1.08350	2.7	60.9
14	1.06256	2.7	63.6
15	1.03362	2.6	66.2
16	1.01867	2.5	68.7
17	.98501	2.5	71.2
18	.97855	2.4	73.6
19	.96689	2.4	76.0
20	.95609	2.4	78.4
21	.90459	2.3	80.7
22	.89136	2.2	82.9
23	.85423	2.1	85.0
24	.78502	2.0	87.0
25	.77952	1.9	89.0
26	.77480	1.9	90.9
27	.72769	1.8	92.7
28	.67232	1.7	94.4
29	.60862	1.5	95.9
30	.51051	1.3	97.2
31	.31564	.8	98.0
32	.24648	.6	98.6
33	.23578	.6	99.2
34	.14252	.4	99.5
35	.06552	.2	99.7
36	.05061	.1	99.8
37	.04827	.1	99.9
38	.01360	.0	100.0
39	.00728	.0	100.0
40	.00000	.0	100.0

TABLE 3.4 - COMMUNALITIES

Variable	Communality
AGE1	.75805
AGE2	.83809
AGE3	.88442
AGE4	.75030
SEX1	.58741
CSE	.27412
OCC1	.58507
DSURF	.82835
DBEACH	.94608
DPARK	.89265
ADJINC	.54023
ANZ	.54981
MCHLD	.35876
NBABY	.46522
NA	.46624
EMP	.46613
FREQS1	.77601
FREQS2	.80836
FREQS3	.59429
FREQS4	.67980
FREQS5	.64166
FREQS6	.41528
FREQS7	.61471
FREQS8	.54478
FREQS9	.39444
FREQS10	.64630
FREQS11	.82137
FREQS12	.65404
FREQS13	.86744
FREQS14	.51290
FREQS15	.97492
FREQS16	.96958
FREQS17	.96334
FREQS18	.96677
FREQS19	.97018
FREQS20	.55761
FREQS21	.89278
FREQS22	.60529
FREQS23	.52866
FREQS24	.88767

(i) football spectating (p = .98)[1]
(ii) strolling in bush (p = .94)
(iii) hunting (p = .73)
(iv) unspecific rural outings (p = .73)
(v) visiting beach (p = .67)
(vi) vehicular bush activities (p = .58)
(vii) swimming (p = .37)

Significant negative loadings are:

(viii) bank/pier fishing (p = -.15)
(ix) water skiing (p = -.14)
(x) picnicking and barbequing (p = -.10)

The significant socio-economic characteristics are, in order:

(xi) passively employed (p = .08)
(x) completed secondary education (p = .07)
(xi) female (p = .06)
(x) income (p = .05)

The last four are characterised by only slight significance.

The dominant behaviour patterns in Factor 1 appear to be low energy ((i), (ii), (iv), (v)), non participatory spectating ((i), (ii), (v)), frequently in a coastal ((v), (vii)) or bushland ((ii), (iii), (iv), (vi)) setting. It is significant that visiting a beach dominates swimming at a beach. There also appears to be something of a 'larrikin element' to it ((i), (iii), (vi)).

This factor appears to be a genuine psychological dimension for it is not heavily correlated with any socio-economic character-

---

(1) We shall employ p to stand for the loading of a variable in a factor. Significance levels may be obtained from Fig. 3.3.

istic - Factor 1 ranges across all classes, sexes and incomes (in fact what correlations there are with Income etc. may be merely a mobility factor). We shall call it the LARRIKIN-TOURIST factor. The inclusion of an uncommon pursuit like HUNTING may disturb the reader and he should be careful to eschew statements like 'The LARRIKIN-TOURIST spends time shooting rabbits. Rather the casuality is likely to be the opposite way ... 'The rabbit-shooter has a LARRIKIN-TOURIST facet to his behaviour.')

Factor 2: This factor may be disregarded as its function is the technical one of removing the bi-polar geographic nature of the observations - some from Melbourne, some from Geelong. We shall call it LOCATION.

Factor 3: The most significant behavioural loadings in Factor 3 are:

(i) site seeing (p = .98)
(ii) picnicking (p = .96)
(iii) hiking (p = .08)
(iv) bank/pier fishing (p = .08)
(v) swimming (p = .07).

Negative correlation are:

(i) boat fishing (p = -.06)
(vii) surfboard riding (p = -.04).

The significant socio-economic loadings are:

(viii) female (p = .05)
(ix) passively employed (p = .05).

The "flavour" of Factor 3 seems to be non-coastal ((iii), (iv), (vi), (vii)), and non participatory ((i), (ii), (iii), (iv), (v)) but not low energy ((iii), (v)): in particular note that swimming is a significant loading but that visiting a beach is not.

This dimension appears to be more purposeful than the TOURIST factor - unspecific rural activities have been resolved into site-seeing or picnicking. Once again the absence of strong socio-economic correlations leads to the feeling that this is a genuine psychological dimensions. We shall call this the NATURE-LOVER factor.

The LARRIKIN-TOURIST and the NATURE-LOVER identify two dimensions of visiting. The former corresponds to spectating and unstructured rural mobility. The latter is also essentially non participatory, but corresponds to a more structured outlook - wherein specific purpose activities (sight-seeing) are preferred to general activities (unspecific rural). The NATURE-LOVER factor indulges in higher energy activities (iii) but these are usually non coastal.

<u>Factor 4</u>: The most significant behavioural loadings in this factor are:

(i) boat cruising (p = .15)
(ii) surfboard riding (p = .09).

A significant negative loading is:

(iii) hiking (p = -.13).

In the socio-economic section the significant positive loadings are:

(iv) 20-35 Age group (p = .90)
(v) number of babies (p = .62)
(vi) completed secondary education (p = .25).

and the negative loadings are:

(vii) 35-60 Age group (p = -.57)
(viii) 60+ Age group (p = -.10).

The clue to this factor is its correlation with child bearing years and number of infants. This dimension is the YOUNG-MARRIEDS and this name will be adopted. The positive correlation with surfboard riding is due no doubt, to residual pre-marriage behaviour; the correlation with boat-cruising is harder to understand. Examination of the correlation matrix casts a little light on this phenomenon, however. In fact boat cruising has, of all activities, the fifth highest correlation with number of babies (below swimming and visiting at beaches, outdoor sporting activities and football spectating). It would seem that boat cruising can be an escape from infant-induced immobility. Indeed YOUNG-MARRIEDS is characterised by a low level of recreation: 14 of 24 frequencies have negative loadings; another 8 are insignificant at the 1% level. This draws attention to what is, one feels, a significant social problem, especially since it occurs in what could be the most active part of the life cycle.

Unlike LARRIKIN-TOURIST and NATURE-LOVER, YOUNG-MARRIEDS is not a psychological dimension, but a socio-economic one - that is, it does not describe a propensity for a certain type of behaviour but rather a certain stage in the life cycle.

Factor 5: The most significant behavioural loadings on Factor 5 are:

(i) surfboard riding ($p = .25$)
(ii) sailing ($p = .18$)
(iii) swimming (protected beach) ($p = .18$)
(iv) swimming (fresh water) ($p = .15$)
(v) rowing and canoeing ($p = .13$)
(vi) surf swimming ($p = .07$).

There is one significant negative behavioural loading:

(vii) golf ($p = -.08$).

TABLE 3.5 - OBLIMAX FACTOR ANALYSIS (AFTER ROTATION WITH KAISER NORMALIZATION)

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
AGE1	.00939	-.00728	-.03101	-.26074	.80386	.00768	.13430	.02894	.03003	.06535
AGE2	-.01230	-.00269	.02352	.89621	.00824	-.06303	.07090	.02282	.01131	-.04220
AGE3	.00766	.03975	.02470	-.56929	-.61106	.03728	.46540	-0.3290	-0.3576	-.02526
AGE4	-.00457	-.04089	-.02886	-.10086	-.06316	.02200	-.83968	-.01620	.00079	.01659
SEX1	.05750	.04761	.05270	.08387	-.02847	.05324	.00027	.03952	-.09617	-.12675
CSE	.06773	.13538	.00100	.25030	-.06740	.04430	.00253	-.12017	.17250	.01271
OCC1	.08362	-.01456	.05235	.06188	-.08404	.04001	.15408	-.00136	-.17512	-.09050
DSURF	-.01634	.89292	-.02180	-.04874	-.01305	-.00556	.00117	.01420	-.05120	.00938
DBEACH	-.00872	.97435	-.01324	.01567	.01895	-.00674	.00762	.00638	-.01668	.00654
DPARK	.01101	.94363	.02063	-.00402	.00584	-.00789	.01206	-.00058	.01782	.00126
ADJINC	.05445	-.02607	-.03582	.11103	-.15731	.02073	.14046	-.03098	-.00119	.04839
ANZ	.03464	.03336	.04703	.04702	.45780	-.01476	.04904	-.05726	-.06278	-.12363
NCHLD	-.01412	-.01689	-.03806	-.00340	-.13924	.02116	.38206	.00863	.02432	.15268
NBABY	-.02779	-.03522	-.02017	.626546	-.15347	.02875	.10112	.06701	-.03095	.02455
NA	.01932	.01396	-.01929	-.05829	-.10837	.03330	-.55732	.01594	-.02826	.04890
EMP	-.00247	.00433	.02531	-.02374	.21745	.00431	.10171	.05866	-.07285	-.02915
FREQS1	.66596	-.02106	.02313	-.02021	-.02458	-.40302	.01014	-.03334	.01822	.02551
FREQS2	.37418	-.01764	.06710	.03809	.07018	.04777	.00094	-.02877	-.02984	.75459
FREQS3	-.02795	-.00095	.02693	-.04748	.17559	-.00151	-.04646	-.04901	-.05084	-.04107
FREQS4	.02944	.02170	.02445	-.02146	.15066	-.01183	.04670	.03485	.04360	-.10173
FREQS5	.01223	-.00760	-.01267	.06201	.12760	.00944	.05213	-.04458	.01411	.00482
FREQS6	.00137	-.02907	-.04027	.09230	.25315	.02150	.11650	-.06809	.11209	.08292
FREQS7	-.14490	.01138	-.01304	.03016	.04856	-.77569	.04072	-.12368	.00561	.03194
FREQS8	.04035	-.02833	-.00239	.15089	-.02363	.00286	.00788	-.39223	-.07130	-.01425
FREQS9	-.01050	.01180	-.00815	-.02813	.18472	-.00111	.01875	-.00296	.03943	.08453
FREQS10	-.03503	-.01447	-.05679	.00516	-.01740	-.09264	.04184	-.78505	-.00616	.06707
FREQS11	-.14543	.02375	.07575	-.02862	-.00055	-.05078	-.04655	.01972	-.05321	.90392
FREQS12	.00132	-.00673	.00891	-.00907	-.00612	-.00055	.05193	-.00580	.80168	-.02370
FREWS13	.94518	-.01344	.01794	-.02899	.01668	.07652	-.00300	-.01410	.01470	.02078
FREWS14	.02452	.00605	.07749	-.12542	-.00656	.03452	-.05967	-.66387	.00163	-.07008
FREQS15	.73003	.00066	.04110	-.00441	-.01669	-.47304	-.00960	.04727	-.00281	.01651
FREQS16	.58109	.00047	.04667	-.00028	-.02573	-.63243	-.00918	.05751	-.00433	.00987
FREWS17	.05894	-.01008	.96651	.00591	-.02520	-.04685	.01269	.01378	02064	-.00445
FREQS18	-.10568	-.00469	.98624	.01345	-.00929	.05714	.01199	-.01990	.02465	.08481
FREQS19	.72949	-.00164	.03859	-.00555	-.02010	-.47231	-.00354	.04393	-.00435	.01036
FREQS20	-.01614	.05725	-.02622	-.01531	-.07988	-.02512	-.08123	.08072	.09009	.04962
FREQS21	.09134	.00215	.05912	-.00165	-.05286	-.90268	-.01510	.05664	-.00463	.00021
FREQS22	.00553	-.03677	.02151	-.01985	.00333	-.00077	-.03270	.02941	.75683	-.04489
FREQS23	-.00366	-.01026	-.02348	.03694	-.01110	.00399	.01515	-.09768	-.01983	.02915
FREQS24	.97523	.01517	.00561	.01307	.02444	.27754	-.00768	-.00910	-.00708	.03657

There are three significant positive socio-economic loadings:

(viii) 13-20 ages (p = .80)
(ix) Australian (p = .46)
(x) Unemployed (p = .22).

Significant negative loadings are:

(xi) 35-60 Age
(xii) Income
(xiii) number of babies
(xiv) number of children
(xv) not passively employed.

This factor is quite unambiguous: young (viii) Australians (ix) involved in water sports (i) (ii), (iii), (iv), (v), (vi). We shall call it YOUNG-SURF but the reader should not read too much into this as the full range of water sports is represented. This factor includes a slight but widespread aversion to non-beach activities - it will be observed that with the exceptions of bush strolling and outdoor sporting activities YOUNG-SURF has a small negative correlation with all activities from FREQS10 (Boat fishing) to FREQS23 (Outdoor sport spectating). Worth comment is the slight (but positive) influence of football spectating and the tiny influence of outdoor sporting activities.

YOUNG-SURF differs from the proceeding factors in that is is mixed - it has both a behavioural and a socio-economic dimension.

Factor 6: The two significant behavioural loadings in Factor 6 are:

(i) football spectating (p = .28)
(ii) strolling in bush (p = .07).

The negative loadings are:

(iii) visiting historical centres (p = -.90)
(iv) Water-skiing (p = -.77)
(v) vehicular bush activities (p = -.63)
(vi) Hunting (p = -.47)
(vii) unspecific rural outings (p = -.47)
(viii) visiting beach (p = -.40)
(ix) boat, sirf fishing (p = -.09).

There are no SES loadings, positive or negative, that are significant at the 1% level. At the 5% level there is a significant negative influence exerted by the 20-35 Age group and females just creeps in.

The dominant characteristic of this category appears to be football going allied with a violent antipathy to all other recreational activities. There is a particular tendency to Philistinism ((iii)) and against activities that require a high degree of organisation ((iv), (v)); the bush is adjured ((vi), (vii)). ((ii)) is against this interpretation but is just on the border line of significance - moreover it is surely significant that it is the least involving of all bush activities); nor is the coast preferred ((viii).)

This factor occurs across all SES groups and seems to be a genuine behavioural dimension - we shall call it FOOTBALL-GOTHIC.

One feels that it would not be self indulgent to call this an unlikeable pattern of behaviour - it is not redeemed, for example by the busyness implicit in the larrikin aspect of Factor 1. Observe that car availability is less than average, although not statistically significant (p = 0.033). The suggestion is that part of this behaviour could be due to enforced immobility. Against this interpretation we have slight upward scores on income and education.

Factor 7: The two significant behavioural loadings in factor 7 are:

(i) sailing (p = .12)
(ii) surfboard riding (p = .12).

The negative loadings are:

(iii) golf (p = -.08)
(iv) hiking (p = -.06).

The positive socio-economic loadings are:

(v) 35-60 Age (p = .47)
(vi) children (p = .38)
(vii) babies (p = .10)
(viii) passively employed (p = .15)
(ix) 13-20 age (p = .13).

Significant negative loadings are:

(x) 60+ Age (p = -.56)
(xi) no access to car (p = -.56).

This dimension describes participation in water sports ((i), (ii)) that passes with the onset of old age ((v), (vi), (vii), (ix), and particularly (x)). This factor is in addition to Factor 5. As a given individual ages, his score on the YOUNG-SURF dimension will decline. It would seem, however, that Factor 7 is needed to explain an additional decline in water sport activities as the individual becomes increasingly infirm. We shall call this factor YOUNG-FIT.

Factor 8: The significant behavioural loading is:

(i) golf (p = .08).

Negative loadings are:

(ii) boat fishing (p = -.74)
(iii) hiking (p = -.66)
(iv) boat cruising (p = -.39)
(v) water skiing (p = -.12)
(vi) outdoor sport recreation (p = -.10)
(vii) surfboard riding (p = -.09).

The significant SES loading is:

(viii) number of babies (p = .07).

and the negative loading is:

(ix) education (p = -.12).

This factor and the following three are quite speculative and we do not wish to claim much importance for them. Factor 8 appears to describe a disinclination to participate in water sports ((ii), (iv), (v), (vii) and note also that 9 out of 10 water activities have negative correlations). We shall call this factor ANTI-WATER.

<u>Factor 9</u>: The significant behavioural loadings are:

(i) skin diving (p = .80)
(ii) outdoor sporting activities (p = .76)
(iii) golf (p = .09).

The negative loading is:

(i) boat cruising (p = -.07).

The significant SES loading is:

(ii) education (p = .17).

and the negative loadings are:

(iii) passively employed (p = -.17)
(iv) male (p = -.10).

This is an interesting dimension: highly active ((i), (ii)) educated ((ii)), males who are not passively employed ((iii), (iv)) are correlated. We shall call this factor SUPERIOR-ACTIVE.

Factor 10: The most significant behavioural loadings are:

(i) bank/pier fishing (p = .9)
(ii) swimming surf beach (p = .75)
(iii) picnicking (p = .08)
(iv) sailing (p = .08)
(v) surfboard riding (p = .08).

The significant negative loading is:

(vi) hiking (p = -.07).

The significant SES loading is:

(vii) 13-20 Age (p = -.12)

and the negative loadings are:

(viii) female (p = -.12)
(ix) Australian (p = -.12)
(x) passively employed (p = -.00).

This factor does not seem to have a recognisiable interpretation. The reader is invited to exercise his ingenuity, but for want of something better we shall call it BANK-FISHING.

Factor 11: The most significant behavioural loadings are:

(i) boat cruising (p = .09)
(ii) outdoor sporting activities (p = .07)
(iii) sailing (p = .07).

The negative loadings are:

(iv) golf (p = -.37)
(v) surfboard riding (p = -.37).

The positive SES loadings are:

(vi) female (p = .68)
(vii) children (p = .21)
(viii) income (p = .14)
(ix) babies (p = .12)
(x) Australian (p = .08).

Negative loadings are:

(xi) passively employed (p = -.67)
(xii) 20-35 Age (p = -.07).

This factor is rather more suggestive than the proceeding one, it seems to be related to femaleness ((vi)) in particular the middleaged ((xii), (vii), (v)) and well off ((vii)) females. There is a strong behavioural content in this factor but those activities that do scrape in have a middle class flavour; we shall call this factor BOURGEOISE.

<u>Other Methods</u>

In the foregoing analysis we have isolated and discussed eleven factors associated with recreational behaviour. These are set out in Table 3.6.

TABLE 3.6 - FACTORS

Factor	Name	Cumulative Variance	Varimax Factor
1	LARRIKIN-TOURIST	14.5	w
2	LOCATION	21.2	2
3	NATURE-LOVER	26.3	3
4	YOUNG-MARRIEDS	31.0	4
5	YOUNG-SURF	35.1	6
6	FOOTBALL-GOTHIC	39.1	(-5)
7	YOUNG-FIT	42.9	(-7)
8 (1)	ANTI-WATER	46.2	?
9 (1)	SUPERIOR-ACTIVE	49.5	8
10 (1)	BANK-FISHING	52.2	9
11 (1)	BOURGEOISE	55.4	11

(1) These factors are quite speculative.

There are a number of approaches to the extraction of factors and an important test of the association between the mathematical expression of a factor and the hoped for psychological dimension is the invariance of the factor with the method. One form of factor analysis is VARIMAX, which produces perfectly uncorrelated factors. VARIMAX factor analysis of our data is set out in Table 3.7. We shall not subject this to exhausitive analysis here but the correspondence between our factors and the VARIMAX factors is quite strong. (see Table 3.3). It will be observed that the first seven factors coincide exactly (with two reversals of direction, ie. replacing a factor by its opposite). Note that the SUPERIOR-ACTIVE factor has survived, even advanced in rank. ANTI-WATER appears to have been resolved into other factors. These results seem to comfirm the overall robustness of the analysis.

Conclusion

One great advantage of a factor analysis is that it promotes research empathy with this data. Having performed this analysis only after the main stream of modelling was concluded we found that a number of valuable insights could still be obtained. A large part of the Geelong Recreational Study was to develop methodologies for recreational planning. We would argue that such studies should begin with a Factor Analysis along the lines set out above.

Beyond this consideration, the conclusion to be drawn from the analysis is not our seven archetypes of recreational behaviour (although it will be of considerable interest to see how persistent these are) but the insight that recreational behaviour is caused by the interaction of two streams of influence, psychological and socio-economic. Some factors in recreational behaviour are psychological (FOOTBALL-GOTHIC), some are sociological (YOUNG-MARRIEDS), and some a blend of both (YOUNG-SURF). The existence of psychological dimensions in recreational behaviour implies that models dependent only on socio-economic data are bound to have low explanatory power and if these are to be improved the psychological dimension must be plumed, especially if policy initiatives are directed at psychological rather than social or economic components of behaviour.

TABLE 3.7 - VARIMAX FACTOR ANALYSIS

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
AGE1	-.01130	-.01250	-.02815	-.29355	.00797	.76847	13666	03992	08180	-.02849
AGE2	.03157	-.00256	.01220	.89416	.06135	.03844	.08447	.03681	-.04290	.00519
AGE3	-.01265	.04335	.03277	-.52668	-.04428	-.62551	.44696	-.06816	-.01205	.01989
AGE4	-.01099	-.04024	-.02829	-.11925	-.02732	-.04458	-.83505	.00077	-.01725	-.02231
SEX1	.03173	.03222	.04450	.05552	-.06843	-.00752	-.00854	-.12423	-.11602	-.05484
CSE	.06266	.14812	-.00044	.27594	-.05106	-.07003	.01366	.19175	.01403	.13087
OCC1	.05561	.00395	.05823	.07581	-.03458	-.06250	.15285	-.16938	-.09500	.02327
DSURF	-.02321	.89304	-.01946	-.04615	.00283	-.02823	.00172	-.04869	.00252	-.00911
DBEACH	-.00940	.97046	-.01254	.02098	.00220	.01114	.01379	-.01102	-.00091	-.00292
DPARK	.01864	.93836	.02163	.00534	.00105	.00638	.01899	.02267	-.00456	.00350
ADJINC	.07057	-.03340	-.03502	.13186	-.02799	-.16771	.15252	.01101	.04586	.01570
ANZ	.02748	.03568	.04643	.02237	.01779	.49672	.06605	-.05976	-.12497	.05729
NCHLD	-.01268	-.01137	-.02975	.01146	-.02321	-.16920	.39304	.02622	.16504	-.01097
NBABY	-.02329	-.03456	-.02802	.62858	-.02776	-.14437	.11554	-.01447	.02085	-.95696
NA	.00036	.00569	-.02382	-.07490	-.03970	-.12574	-.56558	-.03472	.03813	-.02544
EMP	-.00296	-.00383	.01510	-.02952	-.00987	.17271	.08825	-.06088	-.01574	-.05196
FREQS1	.83315	-.01917	.04721	-.00431	.26990	-.01628	.01539	.01705	.04914	.04827
FREQS2	.43898	-.02581	.11902	.03810	-.07940	.07127	.05117	-.00377	.75886	.03493
FREQS3	.00000	-.00234	.02202	-.05184	.00868	.13792	-.04669	-.03236	-.04222	.03210
FREQS4	-.00095	.02733	.01265	-.02264	.02776	.10434	.05256	.03870	.00961	-.02884
FREQS5	.00576	.02187	-.01260	.04971	-.01044	.07198	.05486	.02306	.01731	.04674
FREQS6	-.00626	-.04607	-.03789	.10312	-.01147	.27443	.12164	.12721	.08418	.09818
FREQS7	.17344	.00761	-.00080	.02812	.73483	.06118	.05024	.01070	.02946	.13778
FREQS8	.00739	-.04121	-.00573	.15792	.01464	-.02378	.00486	-.07865	-.00055	.40123
FREQS9	-.00885	.00818	.00449	-.04823	.01214	.21684	.12362	.02787	.08766	.00292
FREQS10	-.00916	-.01309	-.04226	.03506	.11240	-.00428	.06078	.00397	.06586	.78202
FREQS11	-.02125	.01597	.12449	-.03963	.08766	-.01499	.00159	-.02172	.88864	-.01957
FREQS12	.00082	-.00688	-.00346	.03159	.00464	.01753	.05131	.79120	.00079	.01501
FREQS13	.90457	-.00825	.04088	-.00685	-.20460	.02280	.00044	.01129	.05173	.02659
FREQS14	.01290	.01487	.08422	-.09952	-.02160	.01475	-.04768	.00530	-.06848	.65612
FREQS15	.92911	.00184	.06626	.00792	.32186	-.00750	-.00580	-.00403	.04163	-.02905
FREQS16	.84980	-.00000	.07000	.00754	.47818	-.01665	-.00582	-.00508	.03109	-.03918
FREQS17	.22213	-.00788	.95089	-.01486	.06417	-.01601	.01642	-.00308	.06612	.00495
FREQS18	.02824	-.00323	.97036	-.01023	-.00287	-.00052	.02264	.00503	.13965	.03492
FREQS19	.92735	-.00031	.06361	.00679	.32095	-.01080	-.00027	-.00607	.03516	-.02591
FREQS20	-.01765	.06675	-.02500	.01015	.04049	-.06683	-.05827	.09983	.03391	-.07661
FREQS21	.48159	-.00186	.07402	-.00584	.80228	-.04528	-.01266	-.00416	.00732	-.04202
FREQS22	.00064	-.03128	.00644	.01683	.00606	.02171	-.02434	.74597	-.02401	-.02695
FREQS23	-.00342	.01004	-.01780	-.04838	-.01296	-.00578	.03865	-.00667	.01871	.09611
FREQS24	.85018	.02109	.02676	.03594	-.39100	.02846	-.00277	-.00830	.06590	.01970

TABLE 3.7 - VARIMAX FACTOR ANALYSIS

	AGE1	AGE2	AGE3	AGE4	SEX1	CSE	OCC1	DSURF	DBEACH	DPARK
AGE1	1.00000	-.31152	-.35377	-.20731	-.03426	-.06723	-.07998	.01405	-01451	-.02441
AGE2	-.31152	1.00000	-.48657	-.28513	-.00379	.17440	.12151	-.02891	.00575	-.01101
AGE3	-.35377	-.48657	1.00000	-.32380	.00489	-.05854	.03175	.06190	.02984	.04223
AGE4	-.20731	-.28513	-.32380	1.00000	.03465	-.07048	-.10884	-.05966	-.03085	-.01557
SEX1	-.03426	-.00379	.00489	.03465	1.00000	-.07088	-.16538	-.01202	.02508	.01677
CSE	-.06723	.17440	-.05854	-.07048	-.07088	1.00000	.01260	.08750	.08592	.09387
OCC1	-.07998	.12151	.03175	-.10884	-.16538	.01260	1.00000	.07540	-.01159	-.00698
DSURF	.01405	-.02891	.06190	-.05966	.01202	.08750	.07540	1.00000	.81446	.71621
DBEACH	-.01451	.00575	.02984	-.03085	.02508	.08292	-.01159	.81446	1.00000	.91727
DPARK	-.02441	-.01101	.04223	-.01557	.01677	.09387	-.00698	.71621	.91727	1.00000
ADJINC	-.06465	.05685	.06421	-.08660	.01722	.02448	-.02417	-.10900	-.00746	.02789
ANZ	.13809	.01368	-.12299	-.00265	.05811	-.03243	-.00312	-.04661	.05847	.07848
NCHLD	-.01314	-.02805	.12842	-.11890	.01311	-.01401	-.01039	-.00976	.02173	.01735
NBABY	-.14606	.33120	.11876	-.10441	-.00949	.03923	.00818	-.03703	-.00508	-.01895
NA	-.03823	-.10092	-.05751	.24204	.07176	-.06049	-.09745	.03822	-.00991	-.04302
EMP	.10966	-.00439	-.03759	-.06157	-.05068	-.01286	-.05793	-.01809	-.00992	-.00825
FREQS1	.00138	.03075	-.01158	-.03270	.00284	.03020	.02652	-.03629	-.02778	.00756
FREQS2	.06227	.03455	-.02551	-.07587	-.00233	.02419	-.00172	-.04837	-.02640	-.01079
FREQS3	.03397	-.00891	-.02220	.00419	-.01464	-.00551	-.01611	-.01722	-.01469	-.00466
FREQS4	.11438	-.01236	-.01546	-.08552	-.01741	.00405	-.03811	.06955	.05182	-.01686
FREQS5	.02696	.01980	-.01691	-.03125	-.04217	.03534	.01021	.04226	.05143	.01589
FREQS6	.09800	.08421	-.10187	-.07604	-.09447	.03354	-.01510	-.10931	-.03513	-.00718
FREQS7	.02156	.07961	-.04241	-.06702	-.04064	.03922	.04061	.00064	.01695	-.00756
FREQS8	-.02117	.08314	-.04083	-.02822	-.00102	-.00039	.03304	-.04385	-.02311	-.02721
FREQS9	.04980	-.01538	-.01824	-.00968	.01752	-.00412	-.01148	.01279	-.00066	.00321
FREQS10	-.00040	.04701	-.00492	-.06315	-.08622	.02515	.03505	-.00678	-.01252	-.00628
FREQS11	.04999	-.04344	.04430	-.05654	-.06501	-.02219	-.00787	.03262	.00300	.00075
FREQS12	.03555	.07339	-.06608	-.04312	-.07505	.01708	-.00351	-.03350	-.01867	-.00303
FREQS13	-.00473	.02302	-.01817	-.00004	.01808	.06966	.04312	-.03187	-.01819	.01090
FREQS14	.01033	-.03301	.01528	.01039	-.00985	.04385	.05824	.02741	-.00222	.00055
FREQS15	-.01265	.05354	-.02797	-.01702	-.00655	.03937	.03792	-.01769	-.00954	.02030
FREQS16	-.01123	.05316	-.02826	-.01767	-.00934	.03017	.02630	-.02020	-.01020	.02046
FREQS17	-.01210	-.00381	.03811	-.03224	.00661	.00308	.04609	-.02772	-.02232	.02111
FREQS18	-.00220	-.01108	.04295	-.03907	.00228	.00569	.03219	-.02230	-.01093	.01698
FREQS19	-.01539	.04970	-.02141	-.01789	-.00013	.03559	.03684	-.02279	-.01077	.02147
FREQS20	-.04743	.01591	.02100	.00290	-.07076	.05253	.06775	.01817	.05774	.08627
FREQS21	-.01488	.04228	-.01890	-.01243	-.02232	.00501	-.00425	-.01398	-.00890	.01462
FREQS22	.02348	.04432	-.07161	.01328	-.03841	.02369	-.03945	-.04612	-.02534	.00137
FREQS23	-.00169	.00221	.01789	-.02437	-.01857	.04368	-.00149	.02142	.03518	.02487
FREQS24	-.00679	.03854	-.02355	-.01026	.01090	.05794	.04728	.00191	.01981	.02190

TABLE 3.7 - VARIMAX FACTOR ANALYSIS

	ADJINC	ANZ	NCHLD	NBABY	NA	EMP	FREQS1	FEEQS2	FREQS3	FREQS4
AGE1	-.06465	.13809	-.01314	-.14606	-.03823	.10966	.00138	.06227	.03397	.11438
AGE2	.05686	.01368	-.02805	.33120	-.10092	-.00439	.03705	.03455	-.00891	-.01236
AGE3	.06421	-.12299	.12842	-.11876	-.05751	-.03759	-.01158	-.02551	-.02220	-.01546
AGE4	-.08660	-.00265	-.11890	-.10441	.24204	-.06157	-.03270	-.07587	.00419	-.08552
SEX1	.01722	.05811	.01311	-.00949	.07176	-.05068	.00284	-.00233	-.01464	-.01741
CSE	.02448	-.03243	-.01401	.03923	-.06044	-.01286	.03029	.02419	-.00551	.00405
OCC1	-.02417	-.00312	-.01039	.00818	-.09745	-.05793	.02652	-.00172	-.01611	-.03811
DSURF	-.10900	-.04661	-.00976	-.03703	.03822	-.01809	-.03629	-.04837	-.01722	.06955
DBEACH	-.00746	.05847	.02173	-.00608	-.00991	-.00992	-.02778	-.02640	-.01469	.05182
DPARK	.02789	.07848	.01735	-.01895	-.04302	-.00825	.00756	-.01079	-.00466	-.01686
ADJINC	1.00000	-.02984	.03530	.09878	-.05893	-.04112	.04217	.07612	.07073	.01779
ANZ	-.02984	1.00000	-.02709	-.01182	-.07317	.00810	.00231	-.01089	.01102	.02052
NCHLD	.03530	-.02709	1.00000	.09040	-.09550	.02818	.00250	.05506	-.00642	.06082
NBABY	.09878	-.01182	.09040	1.00000	-.03914	-.02372	-.00477	-.01117	-.01930	.00708
NA	-.05893	-.07317	-.09550	-.03914	1.00000	.01416	-.03127	-.01923	-.01620	.03115
EMP	-.04112	.00810	.02810	-.02372	.01416	1.00000	-.01131	.01195	.00384	.00236
FREQS1	.04217	.00231	.00250	-.00477	-.03127	-.01131	1.00000	.37897	.02402	-.01694
FREQS2	.07612	-.01089	.05506	-.01117	-.01923	.01195	.37897	1.00000	-.00505	-.02330
FREQS3	.07073	.01102	-.00642	-.01930	-.01620	.00384	.02402	-.00505	1.00000	-.00350
FREQS4	.01779	.02052	.06082	.00708	.03115	.00236	-.01694	-.02330	-.00350	1.00000
FREQS5	-.01762	.00305	.03504	-.00117	-.01168	-.01282	.01327	.02364	-.00401	.04047
FREQS6	-.01491	.06397	.00752	.00735	-.03216	.03956	-.01256	.07320	-.00963	-.02088
FREQS7	-.03069	.05151	-.01421	-.03329	-.04777	-.01624	.28852	.06362	.01661	.07889
FREQS8	-.01299	-.02138	-.02217	.02227	.00671	-.00502	.02643	.02378	-.00995	.10436
FREQS9	-.02412	.02386	-.01152	-.01141	-.01311	-.00604	-.00527	.07839	-.00147	-.00735
FREQS10	.01935	.01549	.02942	.01042	-.04502	-.01461	.06780	.04423	.00236	-.01063
FREQS11	-.01469	-.03149	.07027	-.01621	-.01566	-.00514	.05965	.54319	-.00966	.06981
FREQS12	-.02161	-.00270	.01800	.01903	-.01759	-.01399	-.00527	.01162	.00529	-.01628
FREQS13	.04692	.01957	.01536	-.03723	-.01213	-.01592	.66133	.41917	.01930	.01854
FREQS14	-.02216	.02970	-.02907	-.02939	-.01790	-.00909	.03735	-.00695	-.00532	-.01000
FREQS15	.06510	.02198	-.01550	-.01833	-.01360	-.00210	.83226	.41027	.01521	-.00483
FREQS16	.06821	.02062	-.01354	-.01341	-.01280	.00264	.81345	.36165	.01303	-.00856
FREQS17	-.00140	.00683	.01117	-.03115	-.02447	-.00257	.24728	.24554	-.00146	.00124
FREQS18	-.01061	.01815	.01376	-.02324	-.02739	-.00978	.08316	.23825	-.00123	-.00139
FREQS19	.06159	.02187	-.01172	-.01287	-.01537	-.00864	.83173	.41174	.01322	-.00855
FREQS20	.05081	.03519	.01159	.03845	-.04881	-.02246	.02052	.00133	-.00815	.12009
FREQS21	.05904	.00796	-.01040	-.01912	-.01115	-.00519	.80830	.16357	.00387	-.00315
FREQS22	.03922	.03134	-.00380	.02602	-.03696	-.01182	.03066	-.00232	.00061	.08613
FREQS23	-.00942	.05998	.02881	.01274	-.02685	-.00006	-.01069	-.00819	-.01394	-.00037
FREQS24	.06362	.02576	-.00734	.02462	.00738	-.00729	.56432	.41967	.01408	.02020

TABLE 3.7 - VARIMAX FACTOR ANALYSIS

	FREQS5	FREQS6	FREQS7	FREQS8	FREQS9	FREQS10	FREQS11	FREQS12	FREQS13	FREQS14
AGE1	.02696	.09800	.02156	-.02117	.04980	-.00400	.04999	.03555	-.00473	.01033
AGE2	.01980	.08421	.07911	.08314	-.01538	.04701	-.04344	.07339	.02302	-.03301
AGE3	-.01691	-.10187	-.04241	-.04083	-.01824	-.00402	.04430	-.06608	-.01817	.01528
AGE4	-.03125	-.07604	-.06702	-.02822	-.00968	-.05315	-.05654	-.04312	-.0004	.01039
SEX1	-.04217	-.09447	-.04064	-.00102	.01752	-.08622	-.05501	-.07505	.01808	-.00985
CSE	.03534	.03354	.03922	-.00039	-.00412	.02515	-.00219	.07109	.06966	.04385
OCC1	.01021	-.01510	.04061	.03304	-.01148	.03505	-.00787	-.01851	.04312	.05824
DSURF	.04224	-.10031	.00064	-.04385	.01279	-.00678	.03262	-.03350	-.03187	.02741
DBEACH	.05143	-.03513	.01695	-.02311	-.00066	-.01252	.00300	-.01867	-.01819	-.00222
DPARK	.01589	-.00718	-.00756	-.02721	.00321	-.00628	.00075	-.00303	.01090	.00055
ADJINC	-.01762	-.01491	-.03069	-.01299	-.02412	.01935	-.01469	-.02161	.04692	-.02216
ANZ	.00305	.06397	.05151	-.02138	.02386	.01549	-.03149	-.00270	.01957	.02970
NCHLD	.03504	.00752	-.01421	-.02217	-.01152	.02942	.07027	.01800	.01536	-.02907
NBABY	-.00117	.00735	-.03329	.02227	-.01141	.01042	-.01621	.01903	-.03723	-.02939
NA	-.01169	-.03216	-.04777	.00671	-.01311	-.04502	-.01566	-0.1759	-.01213	-.01790
EMP	-.01282	.03956	-.01624	-.00502	-.00604	-.01461	-.00514	-.01399	-.01592	-.00909
FREQS1	.01327	-.01256	.28852	.02643	-.00527	.06780	.05965	-.00527	.66133	.03735
FREQS2	.02364	.07320	.06362	.02378	.07839	.04423	.54319	.01162	.41917	-.00695
FREQS3	-.00401	-.00963	.01661	-.00995	-.00147	.00236	-.00966	.00529	.01930	-.00532
FREQS4	.04047	-.02088	.07889	.10436	-.00735	-.01063	.06981	-.01628	.01854	-.01000
FREQS5	1.00000	-.01312	-.01308	.01155	.00708	.04558	.00754	.00721	-.00276	-.00583
FREQS6	-.01312	1.00000	.02100	.04349	-.00031	.01906	-.01813	.06860	.02292	.01202
FREQS7	-.01308	.02100	1.00000	.04975	.02124	.17871	.06686	.02606	.10750	.01086
FREQS8	.01155	.04349	.04975	1.00000	-.00378	.17337	.01147	-.00484	-.00037	-.00877
FREQS9	.00708	-.00031	.02124	-.00378	1.00000	.02149	-.00843	-.00423	.00409	-.00096
FREQS10	.04558	.01906	.17871	.17337	.02149	1.00000	.01203	.09921	-.00130	.21411
FREQS11	.00754	-.01813	.06686	.01147	-.00843	.01203	1.00000	.00358	.03830	-.00860
FREQS12	.00721	.06860	.02606	-.00484	-.00423	.00921	.00358	1.00000	.00952	-.00019
FREQS13	-.00276	.00292	.10750	-.00037	.00409	-.00130	.03830	.00952	1.00000	.01555
FREQS14	-.00583	.01202	.01086	-.00877	-.00096	.21411	-.00860	-.00019	.01555	1.00000
FREQS15	-.00332	-.00747	.35422	-.00500	-.00189	.00920	.06127	.00062	.76120	-.00264
FREQS16	-.00045	-.00546	.42301	-.00048	-.00178	.01131	.06276	.00034	.64815	-.00268
FREQS17	-.00462	-.01649	.08890	.00709	.02929	.00470	.16049	-.00774	.22961	.03953
FREQS18	-.00296	.00363	.03084	.00114	.03320	.01528	.23228	-.00351	.08216	.06348
FREQS19	.00911	-.00309	.35228	-.00682	-.00416	.01149	.04903	.00372	.75603	.00073
FREQS20	-.00927	.00503	.04388	-.02303	-.00265	.01607	.00450	.03494	-.01435	-.01011
FREQS21	.00093	-.00331	.49219	.00581	-.00291	.01306	.07770	-.00365	.24517	.03132
FREQS22	.00606	.01095	.00439	-.02392	-.00715	-.00149	.00438	.27090	-.00414	-.01017
FREQS23	-.00187	-.00077	-.00925	-.00888	-.00558	.00431	-.00426	-.00879	-.01088	.01504
FREQS24	-.00300	-.00796	-.00810	-.00684	-.00101	-.00400	.03103	-.00492	.83969	-.00332

CHAPTER 4 - FACTOR ANALYSIS AND MODEL BUILDING

INTRODUCTION

In this chapter the usefulness of Factor Analysis in building models of recreational demand will be discussed. The methods outlined will be applied to a model of participation in surfboard riding.

Consider a model of the form:

$$y = f(x)+u \qquad \ldots (4.1)$$

where y is some observed behaviour, f(x) represents the systematic dependence of y on the stimulus x and u is a stochastic error. Let us assume that x is a vector of n independent variables. Let us further assume that a factor analysis of a sample of observations of x has produced m significant or identifiable factors: $f_1, f_2, \ldots, f_m$. Then it is possible to consider in place of (4.1) the specification:

$$y = g(f_1, f_2 \ldots, f_m)+u \qquad \ldots (4.2)$$

In the trivial case when m=n (and f and y are linear), (4.1) and (4.2) are equivalent. In general, however, m will be less than n and the two specifications will be inequivalent - the choice between (4.1) and (4.2) will then be decided by which is most accordant with the economic or sociological theory underlying the model. When plausible explanations can be attached to the factor set $f_1, f_2, \ldots f_m$ it is often the case that (4.2) is superior on precisely these grounds.

We have found Factor Analysis to be of value in two distinct areas:

(i) model specification, and
(ii) model interpretation.

In the first case, Factor Analysis frequently enables a more comprehensive model to be specified. The reason is that multi-collinearity, allied with an understandable desire to obtain robust parameter estimates, prompts the analyst to remove or omit variables which are, in fact, quite crucial. The factor approach meets this difficulty by producing an (automatically) orthogonal set of variables so that the temptation is removed.

In the second case the interpretation of a model is sometimes enhanced by the factor approach. It often happens that the influence of a given variable is obscured in that it advances two distinct but oppositive forms of behaviour. For example, membership of the 20-25 Age Group makes it more likely both that an individual be fit and active, and that he be immobolished by babies and young children. The net effect on mobility may be slight - nevertheless it is extremely instructive to decompose this slight effect into large positive and negative components.

These matters will be discussed more thoroughly henceforth.

MULTI-COLLINEARITY

Multi-collinearity is the bane of model development in the social sciences. The econometrician, plagued by correlations between income and car ownership, between elapsed time and distance, between labour and capital, looks with much envy and not a little condescension on the carefully-designed orthogonal experiments of his brother-researcher in the physical sciences. Yet who would be without it? Its value as a whipping boy is learned very early in one's apprenticeship and more model failures have been blamed on multi-collinearity than on all the departures from the assumptions of classical OLS combined. Alas, peccavi, but without a safety net what feats of statistical daring would have been attempted? One senses that where once was jaunty resolution, all would be circumspect and pusillanimous.

The curious thing is that the absence of multi-collinearity is not a prerequisite for OLS. Consider a linear model of the form:

$$y(t) = \sum_{i=1} \beta_i x_i(t) + u(t) \qquad \ldots (4.3)$$

where $x_i(1), \ldots, x_i(t)$ are observations of a random variable x_i, $y(t)$ refers to observations of y, and $u(t)$ is an unobserved stochastic disturbance. We shall write this in matrix notation:

$$y = X\beta + u, \qquad \ldots (4.4)$$

A number of assumptions may be made:

(i) the disturbances have zero mean: $Eu=0$,

(ii) the disturbances have constant variance and are uncorrelated: $E(uu^t) = \sigma^2 I$

where I is the indentify matrix,

(iii) the x_i are distributed independently of u (but not, necessarily, of each other),

(iv) $u \sim n(0,\sigma)$,

and, subject to various of these, a number of deductions made concerning the least squares estimator of β,

$$\hat{\beta} = (X^tX)^{-1}X^tY \qquad \ldots (4.5)$$

with variance-covariance matrix

$$\sigma^2 E(X^tX)^{-1} \qquad \ldots (4.6)$$

The deductions guarantee in a number of senses the optimality of the estimator $\hat{\beta}$.

Note the presence of the expectation operator E in (4.6). In classical least squares the values $x_i(t)$ are not observations of random variables, but rather non-stochastic real numbers. In this case $E(X^tX)^{-1} = (X^tX)^{-1}$. In most applications in the social sciences, however, this is not the case.

Multi-collinearity is not even easy to define except at its two extremes - when it is severe (one variable is an exact linear combination of others), and when it is entirely absent (the explanatory variables are uncorrelated); in any case it has something to do with intercorrelations in the data, and increases as these become more intense.

It will be observed that none of the above requires any assumptions concerning lack of correlation between the x_i. This is a slight overstatement since the estimator in (4.5) demands at least the inversion of the matrix X^tX which will be possible if and only if there is no perfect linear identity in the data X. If, however, there is an approximate linear identify in the data - one variable x_i is nearly a linear combination of some others - then the determinant of X^tX is small and some matrix entries in (4.6) will become very large. This explosion of the variance-covariance matrix results in extremely wide confidence intervals on parameter estimates. Accordingly, the estimates are considerably unstable, in particular with respect to small increases in the data set. A common phenomenon when two explanators are highly correlated is a flip-flop process whereby first one variable is attributed all explanation and, subsequently, for the addition of a handful more observations, it is abandoned in favour of the correlated variable. This happens frequently in transport models where, say, mode choice is specified as a function of time and cost advantages:

$$M = \alpha\Delta T+\beta\Delta C \qquad \ldots (4.7)$$

The pattern of urban travel is usually such that ΔT and ΔC are intensely intercorrelated and in these circumstances it is often found that one of ΔT and ΔC is very significant and the other not at all. All too frequently, dare we say it, the analyst reaches the conclusion that money cost has no bearing on mode choice and all is explained by time differences, or vice versa. Of course, the conclusion is false, but it is doubly pernicious in that the estimate of the significant coefficient is biased upwards by the increased variance the associated variable appears to be explaining. The situation is illustrated in Figure 4.1 depicting the geometric representation of a regression of random variable y on random variables x_1 and x_2. If the variables are assumed to be standardised, then the regression decomposes y as a sum:

$$y = |OP|x_1+|OQ|x_2+y^*$$

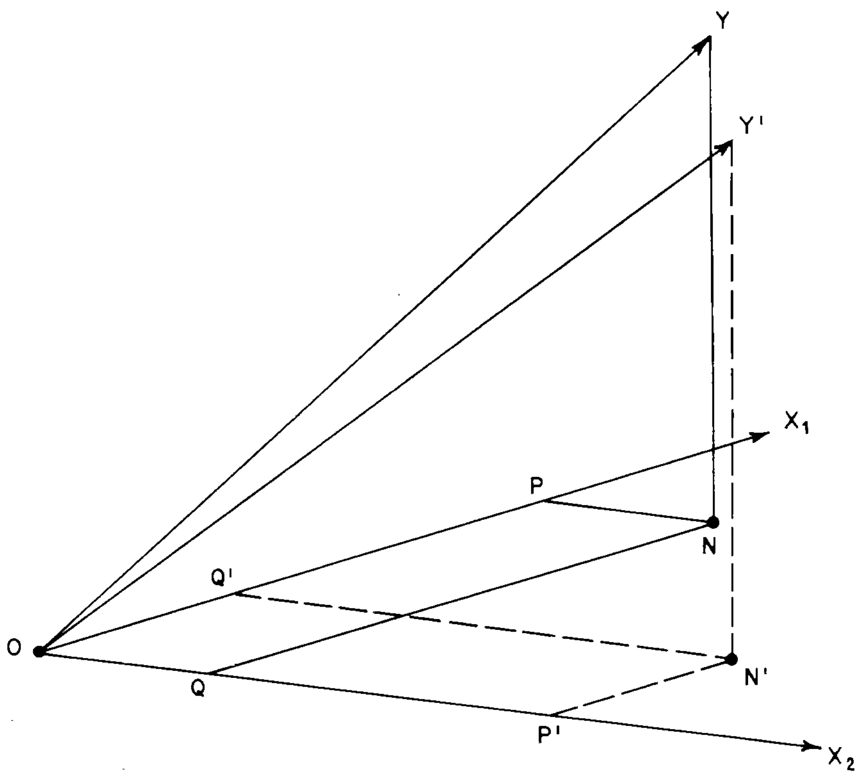
where y^* is uncorrelated with x_1 and x_2.

Note how a relatively minor shift in y(to y') can dramatically reverse the loadings attributed to x_1 and x_2. The reader may care to check that this instability disappears as angle QOQ^1 is increased to 90^o - i.e. as x_1 and x_2 become uncorrelated.

It is important to point out when this can lead to forecasting errors. If, in forecast years, the multivariate distribution of the correlated explanators will remain the same, then nothing is lost - in fact it is perfectly reasonable to omit one or other variable from the specification. On the other hand if the equations are being developed to study the effects of a radical change in this distribution (e.g. in (1.5) faster buses, a new freeway, reduced fares, etc) this mis-specification will lead to biased forecasts in precisely the manner indicated.

This then is the core of the multicollinearity problem, and it is as much psychological as statistical: the analyst is confronted with a choice between a valid specification with theoretically unbiased but completely unbelievable parameter estimates, and a model that is obviously mis-specified but apparently quite robust.

FIG. 4.1 - REGRESSION OF Y ON X_1 AND X_2



MODELLING WITH FACTORS

Multi-collinearity can be quite severe in socio-economic data. An extreme example is the so called 'Dummy variable trap'. Assume the data is naturally divided into four (say) categories and there is reason to believe a dependent variable is separately affected by membership of each of these categories. A model-builder may seek to account for these effects by placing a dummy variable on each category - x_1, x_2, x_3, x_4. But then

$$x_1+x_2+x_3+x_4 = 1$$

an exact liner identify, which means the data is perfectly collinear. The problem may be dealt with by omitting one of the dummies, x_4, say, but if the fourth category is relatively small, or if the remaining variables contribute a fair degree of the explanation of x_4 then the problem returns.

The factor approach to this problem is drastic - the aberrant set of explanators is rejected and replaced by a set of orthogonal (uncorrelated) explanators: that is, the model is re-specified from (4.1) to (4.2). It is not unreasonable to ask what has become of the multi-collinearity. We shall answer this for the 2-dimensional case. Here multi-collinearity is caused by a substantial correlation of the two explanators. Effectively the factor approach argues that this correlation is caused by the presence of a more fundamental entity - the factor - and a factor analysis aims at constructing synthetic measures of these factors from the exhibited pattern of correlations. Thus if there is a strong correlation in the data this is merely evidence of an underlying factor which may be used as a principal explanator. The amount of confidence appropriate to this procedure, is, of course, in direct proportion fo the reasonableness or interpretability of the synthetic factors - only when this is strong may we regard the factor specification as superior to the unfactored specification.

A related plus of this methodology is that the influence of a larger set of variables can be included. For example, even if there is a perfect collinearity in the data, it is still possible to have the offending variable summands in the expressions for the synthetic factors. Frequently this means that the factor specification (4.2) will explain a larger percentage of the variance in the independent variable than the unfactored specification (4.1): the collinearity forces the omission of an important explanator from the latter specification. Usually, however, the factored equation explains a smaller percentage of the variance, as, indeed, might be deduced from the optimal properties of least squares.

The points discussed above relate to model specification. The most notable advantage of the factor approach, however, comes from improved model interpretation. The reason for this is discussed in the introduction to this chapter - frequently the influence of socio-economic descriptors such as Age Group and Income can have equivocal effects on a market-wide demand but dramatic and opposite effects on two submarkets. This will be exemplified in the next section.

THE GEELONG-SURFBOARD RIDING MODEL

The Geelong Recreation Study considered two models related to surfboard riding:

(i) A model of participation

$$P = P(SES)$$

(ii) A frequency model, granted participation

$$F = F(SES)$$

where P is the probability of participation, in surfboard riding, F the frequency of individuals who participate, and SES a vector of socio-economic and locational variables to describe the individual.

It is our intention to compare the results of the first model estimated on factored and unfactored variables. The unfactored results are set out in Table 4.1. We have employed here a linear model of the probability of participation, rather than the Logit form used in the Geelong Study itself. This is, of course, not strictly valid - a linear curve has not the 0, 1 bounds required of a legitimate probability and our original intention was to use the Logit form - but it was felt at the time that the ease and cheapness of a regression fit was sufficient justification. This set of variables was precisely that settled on in the Geelong Study proper and it was decided to employ precisely this set in the present exercise. Note the exclusion of the AGE2, AGE3, and AGE4 variables - this was forced by recurrent multi-collinearity reflected by incorrect signs and large error variances.

The probability model exhibits the pattern of significance set out in Table 4.2.

Notice that the list of insignificant variables includes two that certainly should be strongly negative (NBABY, NCHILD). It seems certain that the multi-collinearity alluded to above has resulted in a mis-specification in the stage in the life cycle variables (Age, babies etc). (Let us observe in passing that this does not seriously affect the model as a forecast tool provided that the multivariate distribution of the collinear variables remains the same.)

TABLE 4.1 - PROBABILITY MODEL

Variable	B	TD Error B	F Significance
SEX 1	-.56491191E-01	.1359412E-01	17.265403 .000
NBABY	-.33213708E-02	.11694876E-01	.80657259E-01 .776
DSURF	-.40696844E-03	.11051963E-03	13.559468 .000
EMP	.88475435E-01	.48391464E-01	3.3427807 .068
ANZ	.41691249E-01	.15515663E-01	7.2202989 .007
NCHLD	.27532530E-02	.72798083-02	.14303831 .705
CSE	.17592812E-01	.21082957E-01	.69631780 .404
ADJINC	-.27419695E-03	.1865061E-03	2.16114191 .142
NA	-.51969756E-01	.24134455E-01	4.6368853 .032
OCC1	-.73225871E-02	.19441574E-01	.14186205 .707
AGE1	.11119449	.17706083E-01	39.438604 .000
(CONSTANT)	.99994730E-01	.18630540E-01	28.807351 .000
MULTIPLE R	.23856		
R SQUARE	.05691		
STD DEVIATION	.28054		
F-STATISTIC	9.74 (11DF)		

TABLE 4.2 - SIGNIFICANCE IN PARTICIPATION MODEL

1%	Z	10%	Insignificant
SEX1 (-1)	NA (-1)	EMP(+)	NBABY(-)
DSURF (-)			NCHILD (+)
ANZ (+)			CSE (+)
AGE1 (+)			OCC1 (-)
			ADJINC (-)

FACTOR ANALYSIS OF THE SES DATA

A Factor Analysis was carried out on all SES variables from the full collection of 1787 individuals. Six factors were selected and the matrix of Factor Scores is set out in Table 4.3. We suggest the following interpretation of factors:

LOCATION: This factor arises from the bipolar nature of the data (half in Melbourne, half in Geelong) and is not particularly instructive.

YOUNG-MATURITY: The strong correlations in this factor are without the 20-35 age group, education and income, and against the <20 age group and the 35-60 age group.

AGED: The negative of Factor 3 is clearly associated with the aged. The strong correlations are with the 60+ age group, and with no car access and against the 35-60 age group, income, and dependent children.

YOUTH: Factor 4 is correlated with the <20 age group and with Australian born, and against the 35-60 age group and no car access. This factor is fairly unambiguously associated with Australian-born youth.

FEMALE: Factor 5 is strongly correlated with female sex, income, and children, and against passive employment and education.

LOWER YOUTH: The last factor is of considerable interest. It is strongly correlated with the <20 age group and unemployment, and negatively correlated with passive employment, Australian born, the 35-60 age group, female sex and income. This factor appears to be associated with lower SES youth.

These factors seem quite sound - our struggles for interpretation were not too outlandish or unseemly. In the next section we shall consider these factors as independent variables in the probability model.

THE FACTORED PROBABILITY MODEL

The results of a regression of P on the six factors discussed in the preceeding section are set out in Table 4.4. It will be observed that all variables are significant at the 5% level, four are significant at the 1% level and three at the .1% level. Moreover all of the parameters have sensible signs, subject to the interpretations we have attached to the corresponding factors. (Note that AGED is the negative of Factor 3 so that the sign on FACT 3 should be reversed.) The variables are in order of importance:

(1) YOUTH
(2) LOWER-YOUTH
(3) AGED
(4) YOUNG-MATURITY
(5) LOCATION
(6) FEMALE

Note that R^2 has slightly increased over the unfactored model, caused by the inclusion of the variables AGE2, AGE3, and AGE4. If we had excluded these variables (strictly speaking only two could have been) the R^2 would of course have shown a decline. In the original analysis these were rejected as they appeared to react deleteriously with the other SES variables.

It must be emphasised that miniscule gains in R^2 are not the object of the exercise - the major aim is to produce a model which is correctly specified - lacking no important variables - and readily interpretable. The factored model appears to achieve these ends.

TABLE 4.3 - FACTOR SCORE MATRIX FOR FULL DATA SET

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
AGE1	-.01471	-.26825	.07004	.66438	-.05355	.50602
AGE2	-.02269	.89409	.06161	.06731	-.09968	-.08695
AGE3	.06143	-.54622	.48694	-.54971	.02383	-.21888
AGE4	-.03634	-.12041	-.78811	-.06951	.15034	-.14122
SEX1	.03540	-.01289	-.02104	.11294	.68232	-.18947
CSE	.14270	.34079	.02419	-.13013	-.22826	.09649
OCC1	.00760	.05881	.12471	.01334	-.66249	-.33337
DSURF	.89681	-.02679	-.02401	-.04084	-.09047	.02584
DBEACH	.97088	.03030	.01404	.02718	.03821	-.00872
DPARK	.93623	.01764	.03200	.03861	.04287	-.03428
ADJINC	-.05194	.15401	.30705	-.08851	.29448	-.16478
ANZ	.04107	-.03329	.07089	.67874	.12524	-.22122
NCHLD	.01075	-.00028	.45022	-.16800	.23417	.17819
NBABY	-.03836	.62315	.16786	-.11146	.18718	-.02882
NA	.00871	-.06932	-.57158	-.26701	.20262	.14772
EMP	-.01011	.03443	.06170	-.06775	-.01826	.74375

TABLE 4.4 - REGRESSION OF P ON SIX FACTORS

Variable	B	STD Error b	F Significance
FACT 1	-.14250159E-01	.65976116E-02	4.6651534 031
FACT 6	.36923383E-01	.65971349E-02	31.325085 .000
FACT 2	.15517642E-01	.65978971E-02	5.5314681 .019
FACT 5	-.25363884E-01	.65975933E-02	14.779524 .000
FACT 4	.51062409E-01	.65971696E-02	59.908341 .000
FACT 3	.20053085E-01	.65972731E-02	9.2391810 .002
(CONSTANT)	.91214770E-01	.65957343E-02	191.25114 .000
MULTIPLE R	25659		
R SQUARE	.06584		
STD DEVIATION	.27882		
F-STATISTIC	20.9 (6DF)		

CHAPTER 5 - CONCLUSION

It has been argued that data collected at a household level is essential if models are to be sensitive to the complete range of recreation opportunities. In the Geelong Recreation Study household information was collected which enables models to be developed to relate socio-economic characteristics with reported recreation behaviour.

The models are suitable for use in predicting changes in level of demand following changes in population socio-economic patterns and site access costs. As we shall demonstrate in the Phase 2 of the Geelong Study the usefulness of the models appear to be considerable. In the second phase a number of possible scenarios of population growth in Geelong and Melbourne will be proposed and used as the base for predicting possible changes in the levels of usage of recreation sites in the Geelong Region.

It should be emphasised that the models have been developed as a tool for predicting future changes in patterns of participation so that planners and park managers will be able to predict where and when the major points of pressure are likely to occur. In this respect the absolute magnitude of visitors at any site may not be critical as it is a knowledge of future changes that is of most use to the planner. If an area is presently under threat of environmental degradation because of over-use, then it is the relative change from the present position that one is interested in more than the present visitor rate. If the present visitor rate to the You Yangs (approximately 200 000 people per year) is considered to be close to the maximum that the park can accommodate without endangering wild life and vegetation in the area, and environmental scientists predict that an increase greater than, say, 5% in the annual visitor rate will cause irreparable harm to the area, then Forest Commission managers will be keen to be advised on means of controlling visits to the area rather than with information on the absolute visits. That is, for planning policy work, relative changes from the status quo are usually more relevant to management decisions than absolute levels of demand.

There are very few opportunities available to compare model estimates with site counts. The most reliable on-site surveys have been undertaken at the You Yangs Forest Park where the total number of annual visitors have been estimated with the aid of a road traffic counter, and the origin of visitors estimated by on-site surveys. A comparison of these results with the model estimates is presented in the following table:

TABLE 5.1 - COMPARISON OF ON-SITE AND MODELLED USAGE INFORMATION FOR YOU YANGS

		On-Site Information	Model Estimate
(i)	Annual Visitors	220 000	340 000
(ii)	Origin of Trips:		
	East of Melbourne	17%	17.1%
	West of Melbourne	30%	32.8%
	Corio)		12.8%)
	Rest of Geelong)	52%	19.0%) 49%
	Elsewhere)		17.1%)
		(N = 555)	

If the on-site information can be believed then it suggests that the model is overestimating the total number of trips being made to the You Yangs by approximately 50%. The prediction of trip origins appears to have been at a surprisingly high level of accuracy. The overestimate could well be explained by the survey response bias discussed in the last section of Chapter 1.

The accuracy with which the origin splits being estimated reflects the sensitivity of Region and Site choice models and indicates that activity location choice will be responsive to changes in population settlement patterns and site access "costs". That is, whilst the Geelong models may overestimate actual trip production rates (perhaps by up to 50%) the models appear to be sensitive in predicting relative changes of usage.

There are two major shortcomings associated with the Geelong Study approach. These shortcomings are discussed below:

(i) The modelling procedure was very expensive, requiring the calibration of 43 individual models. It has been suggested that the number of models can be reduced by imposing a number of simplifying assumptions. The first of these would be to eliminate the Regional choice model ($R_{ij}(g)$ from the sequential process. Such a change is untenable for reasons of both a theoretical and practical nature (see Chapter 1, Formal Presentation of the Models). In any event it would only reduce the number of models to be calibrated by three.

A further simplification could be achieved by combining the first two choice decisions into one; namely calibrating a single model of the number of times an individual will undertake an activity rather than calibrating firstly, X(I), will an individual undertake an activity and then, N(I), if an individual undertakes an activity how often will he do so?

By combining the first two choice processes the total number of models required to be calibrated would be reduced by ten. The cost savings to be gained by introducing such a procedure would not, however, be significant as binary logit modelling is not an expensive process in comparison with the larger multinominal site choice models. In any event, any move to reduce the number of sequential choice processes must be viewed as one of sacrificing the logical "behavioural" nature of the choice process.

Whilst it is not suggested that the overall structure of the Geelong models be changed in any way, it is felt that significant economies could have been achieved by a more prudent specification of the ultimate requirements of the models. That is, it may not have been necessary to develop the full set of models. If, for example, the most important problem facing the Geelong Region is how to control the areas used by sight-seer/picnickers, models of site choice behaviour developed for these activities may only be sufficient to develop sensible management policy. Such models would cost only 10% of the total modelling cost. That is, it would be more economical to develop models as they are needed for management purposes rather than develop the full set of models in one go.

(ii) The second major shortcoming of the models is their insensitivity to site conditions which influence individual site choice behaviour. We know nothing about how the quality of a site affects choice behaviour and we can therefore say nothing about how site choice behaviour will alter as the quality of a site is changed. Such changes in site conditions may be associated with crowding, site development, changes in use, etc. In other words, a glaring omission from the Geelong study is a consideration of how individual attitudes to site and activity attributes affect behaviour.

In respect to the second shortcoming this report has introduced the subject of attitudinal data and explained the way in which it could be incorporated into the conventional Geelong Recreation Study procedure. It should be emphasised that the use of attitudinal data does not represent an alternative methodology but rather an adjunct to the whole study procedure. It should be included to provide further behavioural insight into the choice processes that are being modelled.

Factor analysis has also been discussed as a technique to be used in conjunction with the attitudinal analysis. The strengths of the process have been described in Chapters 3 and 4. When extensive attitudinal data sets are collected with large lists of attributes it is inevitable that intercorrelations among attributes will exist leading to problems in interpreting survey cross-tabulations and model coefficients. Under such conditions it becomes difficult to distinguish between separate influences of the explanatory variables and to obtain estimates of their importance (Recker & Stevens, op.cit., p. 557).

The main thrust of future research into recreation demand modelling should therefore be towards expanding the Geelong Recreation Study procedure to include data on individual attitudes and perceptions of the recreation experience. It is suggested that the techniques developed by Anderson and Louviere may be successfully employed. The series of attitudinal questions suggested in Chapter 2 could be incorporated quite easily into the questionnaire which was used in the Geelong Survey.