An Evaluation of Socio-Economic Classification Systems

A thesis presented in partial fulfilment of the
requirements for the degree of
Master of Business Studies
at Massey University

Robert William Mayo

1985

ABSTRACT

This thesis evaluates some commonly used socio-economic classification systems. Some of the systems evaluated have been used for many years in the market research industry in New Zealand whilst others are recent additions or are more commonly used in the United Kingdom.

The main objective of this study was to test the ability of the systems to predict purchasing levels of consumer products and services. The second objective was to evaluate how well the various systems predict brand choice.

A sample of 1596 repondents was provided by AGB:McNair from their media survey database. Multiple regression was used to predict the level of usage of each product, with the adjusted R^2 value of the equation as the measure of the power of the classification system. Nominal variables, such as brand last used, were crosstabulated against the clasification categories, and Lambdas calculated. A further measure of their ability of the classification systems to predict brand choice was obtained by performing discriminant analysis, which generated classification tables. The percentage of cases correctly classified provided a further measure of performance.

The various classification systems were not very good at predicting purchasing behaviour. The better systems accounted for about 2% or 3% of the variation in quantities purchased. The various classification systems were also not very good at predicting brand choice. Even though the various classification systems explained little of the variation in quantities purchased and brand choices, they are still very useful. The socio-economic classification systems can be used as a starting point from which better predictors of purchasing behaviour can be developed.

ACKNOWLEDGEMENTS

CONTENTS

APPENDICES

BIBLIOGRAPHY

1 INTRODUCTION

1.1 BACKGROUND

This thesis evaluates some commonly used socio-economic classification systems. Some of the systems evaluated have been used for many years in the market research industry in New Zealand whilst others are recent additions or are more commonly used in the United Kingdom.

It is almost universal survey practice to collect personal information about informants and their households, usually at the end of an interview. The personal information collected usually deals with income, occupation, household composition and other demographic details. This personal information is known as classification data, and has a number of uses.

Groups may be formed from the classification data and used to form strata or quotas for surveys, thus ensuring a more representative sample and greater precision. Classification data can provide consistent sub-groups of wide application across many data sources, provided these use a common system of classification. For example, in the United Kingdom buyer behaviour information on one survey can be linked by social grade to media exposure on another. Finally, classification data can be used to locate concentrations of groups of

marketing interest; for example, where the best place is to locate a new supermarket. Some classification systems classify individuals while others classify households. When a household is classified all individuals in the household, regardless of their incomes, age and other details, adopt the household classification.

Many of the classification systems in New Zealand are used because they were successful overseas. There has been no systematic evaluation of the relative merits of the various classification systems to predict purchasing levels of consumer products and services or predict brand choice for New Zealand.

1.2 OBJECTIVES

The main objective of this study is to test the ability of the systems to predict purchasing levels of consumer products and services. The second major objective of this study is to evaluate how well the various systems predict brand choice.

1.3 METHOD

Every week AGB:McNair Limited conduct a regular survey, known as the media survey, of 200 New Zealanders. The survey collects information on magazine readership and purchasing of consumer products and services. The reports constructed from this survey are used by most media houses and advertising agencies in New Zealand. The survey involves personal interviews with respondents over 10 years old from throughout New Zealand. A more detailed description of the survey method is given in Appendix 1. Agreement was obtained from McNairs to use their survey as the database for this study. They agreed to include some additional questions on the demographic composition of the households they interview, thus enabling the construction of the classification systems used in this study. The survey was conducted in early 1984 and included 1596 respondents.

The analysis of the survey was conducted by using the SPSS[x] statistical package. Multiple regression was used to predict the level of usage of each product, with the adjusted R^2 value of the equation as the measure of the power of the classification system.

The adjusted R^2 values were then averaged over product groups to give an indication of the various systems' ability in a specific product group. The adjusted R^2 values for each specific system were also averaged over all products, thus giving an indication of the overall

ability of the various systems

Nominal variables, such as brand last used, were crosstabulated against the classification categories, and lambdas calculated. A further measure of ability of the classification systems to predict brand choice was obtained by performing discriminant analyses, which generated classification tables. The percentage of cases correctly classified provided a further measure of performance.

There are a number of problems in assessing the overall ability of the various classification systems. In this study, all classification systems were tested on all the consumption variables. In practice, some of the systems would not usually be used on some specific products as they are known by common sense to be of little use in prediction of consumption or brand choice. Therefore the inclusion of the R^2s, lambdas and percentage misclassified obtained from these cases has dragged down the overall averages for those systems. Some of the systems are designed to predict individual behaviour while others are designed to predict household behaviour. Much of consumption information used in the evaluation of the systems was on individual consumption, therefore the overall R^2 values are weighted in favour of those systems that were better at predicting individual behaviour. Further details on the statistical analysis undertaken in this study is included in Appendix 1.

2. SUMMARY

Quantities Purchased

The various classification systems were not very good at predicting
purchasing behaviour.   The better systems accounted for about 2% or 3%
of the variation in quantities purchased.   Similar studies in other
countries have also reported only small amounts of the variation in
purchasing behaviour as explained by socio-economic classification
systems. Most of the classification systems predicted the purchase of
personal products, such as shampoo and deodorants, reasonably well.

The following systems performed consistently better than the others:

Sagacity

McNair occupation groups

Age of respondent

Brand Choice

The various classification systems were also not very good at
predicting brand choice.

Various overseas studies provided little evidence to support socio-economic classification systems as good predictors of brand choice.

The following classification systems were the better predictors of brand choice:

Sagacity

McNair occupation groups

Age of respondent

Even though the various classification systems explained little of the variation in quantities purchased and brand choices, they are still very useful. The classification systems can used as a starting point from which better predictors of purchasing behaviour can be developed. This can be done by adding new variables to the classification system. Age of respondent was added to all the classification systems and new estimates of R^2 were calculated for a number of products. The addition of age to the systems increased the amount of variation explained by the systems by about 4%. Many similar adaptions of the standard systems can be performed.

3 DESCRIPTIONS OF CLASSIFICATION SYSTEMS

CLASSIFICATION SYSTEMS USED IN THE UNITED KINGDOM

There are several classification systems widely used in the United Kingdom. The social grade system is universally used by market research agencies. Other systems in use include SAGACITY, ACORN, Leo Burnett Life Style and MONITOR-Value Groups. The Registrar General's systems are used for official statistics.

Social Grade

The most widely used classification system in the United Kingdom is Social Grade. This was developed for the National Readership Survey and is used universally by market research and advertising agencies. The social grade is based on the occupation of the head of the household. Occupation is used as an indication of the status or general standing of the person in the community. The social grade found for the head of the household is applied to all members of the household. If occupational information is incomplete or refused then a number of factors such as qualifications held are collected. See Appendix 2 for details.

The social grades used are shown below.

Grade	Description	% of Households
A	Upper Middle Class	3%
B	Middle Class	13%
C1	Lower Middle Class	22%
C2	Skilled Working Class	32%
D	Semi-skilled and Unskilled Working Class	21%
E	Households at the lowest level of subsistence	9%

A detailed description of the social grades is included in Appendix 2.

Interlaced Demographics

Interlaced demographic classification systems use standard demographic
variables such as age and income of respondent and group the scores on
these variables into a classification system.

A recent development using interlaced demographic data is SAGACITY.
SAGACITY was developed by Cornish (1981) at Research Services Limited.
SAGACITY uses classification data which is routinely collected on the
National Readership Survey and most other surveys.   The aims of
SAGACITY were described by Cornish as "to provide an analysis using the
standard demographic data in a market survey to arrive at a target
market description... to provide a means of assessing media coverage
within the defined target market". The demographic variables used are
life-cycle, income and social grade.   These three variables are
interlaced to provide 12 classification groups.

The life-cycle stages used are shown below.

Life-cycle stage	Description	% of adult population
Dependant:	adults from 15-34 who are not heads of households or housewives, unless they are childless students in fulltime education.	16%
Pre-family:	adults from 15-34 who are heads of households or housewives but are childless.	8%
Family:	adults aged less than 65 who are heads of households or housewives in households containing one or more children under 21.	36%
Late:	includes all other adults whose children have left home or who are 35 years or over and childless.	20%

Social grade groups, "white collar" and "blue collar" are used to split the four life-cycle cells into eight cells. The social grades A, B and C1 are the white collar group and the C2, D and E grades are the blue collar group.

The cells in family and late stages are each split into "better off" and "worse off" groups thus producing four cells which brings the total number of cells to 12.

Income is collected on a ten point scale used on the National

Readership Survey. The National Readership Survey collects the net income of the chief income earner of the household. The SAGACITY system uses this as the basis for the better off and worse off split. Adjustments are made for the working status of the chief income earner's spouse.

Each step on the scale is approximately 25% greater than its predecessor. The scale is adjusted every six months to allow for the predicted rate of inflation over the six month period. With a constant price income scale the real value of the income groups remains constant over time. This scale is used on the National Readership Survey where respondents are asked which one of the income points is closest to the income after taxation of the chief wage earner. The scales used during 1980 are shown below.

Income Code	Income Scales (Sterling Pounds)	
	January-June 1980	July-December 1980
0	1,120	1,210
1	1,400	1,520
2	1,750	1,820
3	2,190	2,370
4	2,730	2,960
5	3,420	3,700
6	4,270	4,630
7	5,340	5,880
8	6,680	7,230
9	8,350	9,040

The cutoff points for better off and worse off vary between the family and late stages of the life-cycle. They also vary between white collar and blue collar occupations.

For example, a white collar adult in the family stage is classified as better off if the head of the household's income is in the two highest of the ten brackets, so long as he or she has no working spouse. If the spouse is working part time, a third income bracket is included; if the spouse is working full time, the four highest brackets are classified as better off.

The following are the 12 groups formed by SAGACITY. The table includes the percentage of households in each group in the United Kingdom.

	Dependant	Prefamily	Family	Late
white collar (A,B,C1)	6.2%	4.0%	better off 6.3%	better off 4.9%
			worse off 7.6%	worse off 9.4%
blue collar (C2,D,E)	9.4%	3.9%	better off 8.7%	better off 6.7%
			worse off 13.8%	worse off 19.0%

Cornish compared the discriminatory ability of SAGACITY with that of age and social grade by comparing the average behaviour of all adults. The comparison used an index format, with 100 representing the average behaviour of all adults. Cornish then pointed out large deviations

from the average behaviour (100) obtained by various systems. The study involved the media exposure of a sample from the National Readership Survey. The age and social grade breakdown suggested that the most intensive audience of independent radio were the blue collar youngest adults. However, the SAGACITY analysis showed that the dependent stage young people in white collar households listened 9% more intensively than their blue collar equivalents. It was only in the prefamily and family stages that blue collar adults emerged as the more intensive listening group.

This extra detail that SAGACITY added also appeared for cinema attendance and Sunday newspapers. SAGACITY also discriminated well for package holidays and basic durables.

Cornish provides evidence that SAGACITY discriminates both in relation to markets and media. SAGACITY adds a great deal of information to the picture of media usage obtained by conventional analysis. SAGACITY often provides discrimination in markets where social grade analysis would provide little discrimination. A comprehensive description of SAGACITY is included in Appendix 3.

Locational Classification

The ACORN system was developed in 1977 at the Centre for Environmental Studies in the United Kingdom. ACORN classifies areas and the households in them by the predominant housing characteristics of the area.

ACORN was developed from a cluster analysis of 40 variables measured on the 1971 census of the United Kingdom. The scores on the 40 variables were computed for a sample of nearly 4000 wards and parishes which are the fundamental district units in the Census. It is unclear which variables were used in the analysis, but Bermingham, Baker and MacDonald (1979) state the following were the 11 "key" variables used:

```
level of unemployment
proportion of students
number of two car households
proportion of immigrants from the Commonwealth
proportion of the population working in a particular
  sector of industry or commerce,
Registrar General's Social Class
age structure
tenure type
five year migrancy
level of overcrowding
level of basic housing amenities
```

These variables aimed to measure the key social, housing and demographic characteristics of the area. A 36-cluster solution was initially attained; this was reduced to a seven-cluster solution.

ACORN was modified at British Market Research Limited   by Bermingham,

Baker and MacDonald (1979).   They felt that clusters 1, 2, 6 and 7

contained very disparate groups. This led them to sub-divide those

groups and produce 11 clusters.   The 11-cluster solution is shown

below.

Group	Title	% of U.K. population
1A	Modern family housing for manual workers	12
1B	Modern family housing for higher incomes	8
2A	Older housing of intermediate status	15
2B	Very poor quality older terraced housing	6
3	Rural areas	6
4	Urban local authority housing	18
5	Housing with most overcrowding	3
6A	Low income areas with immigrants	6
6B	Student and high status non-family areas	5
7A	Traditional high status suburbia	18
7B	Areas of elderly people (often seaside resorts)	4

Bermingham et al. claim that ACORN provides an excellent tool for

sampling and survey design.   ACORN is also claimed to extend social

grade as a measure of status. This claim is based on a comparison of

the discriminatory ability of ACORN and social grade on a study of

heavy wine drinkers.   The social grade analysis showed ABs, who

represent 15 per cent of all adults, account for 49 per cent of heavy

wine drinkers. The analysis using ACORN showed the "jet set" (group

6B) living in the high-status inner city areas, where 35 per cent of

the population are AB, have three and a half times the average wine

consumption. Those in the low income areas with a high immigrant

population (group 6A) are as likely to be heavy wine drinkers as their

better-off counterparts in affluent suburbia (group 7A) or in the

retirement areas (group 7B). The low income areas with a high immigrant population (group 6A) have twice the per capita consumption of those living in the agricultural areas (group 3) where 19 per cent of the population are ABs. The ACORN system has been extensively used during the last year by the UK Post Office to assist direct mail advertisers (Chilvers and Mcmanus 1983). A detailed description of the modified ACORN clusters is included in Appendix 4.