Chapter 4. Methodology

4.1. Introduction

Contemporary literature upon research procedures within a positivist framework[1], generally distinguishes four methods of data collection - observational methods (describing phenomena as they occur in their natural settings), survey research (such as questionnaires and interviews), secondary data analysis (investigating existing documentation on the phenomena under consideration), and qualitative research (such as case studies)[2]

Each of these approaches has its own limitations and advantages and their relevance depends on the subject under investigation and the philosophical approach adopted by the researcher[3]. Given the shortcomings of each method, it is accepted that there may be advantage in using more than one method of data collection to test the same hypothesis in any investigation, the final method or combination of methods eventually chosen in any project depending on the philosophical approach adopted[4], aims of research, the target population and the time and resources at one's disposal.

4.2. Tasks specific to this thesis

Given the thesis put forward above - that Gaelic Language, Arts and Culture has a positive effect on the economic development of the Gaelic economy, given its impact on a number of direct and indirect variables, the tasks facing this thesis were twofold.

On the supply side these included identifying the components of the Gaelic Language Artistic and Cultural sector within the Gaelic economy, aiming to assess its direct and

---

[1] Defined here as one which 'embraces any approach which applies scientific method to human affairs conceived as belonging to a natural order open to objective enquiry' Hollis, M. (1996). The Philosophy of Social Science - an introduction. Cambridge, Cambridge University Press.

[2]Frankfort-Nachmias, C. and D. Nachmias (1992). Research methods in the Social Sciences. Sevenoaks, Hodder & Stoughton.

[3] Whilst positivism (see footnote above), has been adopted as a relevant approach for the aims of this enquiry, it is not the only possible approach to philosophical enquiry. Other philosophical positions may be adopted if the subject matter is concerned with value judgements such the categorisation of qualities of men, events or things as 'noble, ignoble, good, evil, beautiful, ugly etc' Kolakowski, L. (1993). An Overall View of Positivism. Social Research: Philosophy, Politics and Practice. M. Hammersley. London, Sage.. This was not the case with this thesis however.

[4] The approaches adopted within this thesis (particularly survey research and secondary data analysis) are commonly found amongst the research carried out within a positivist framework.

indirect economic impact, and analysing the views held by the individuals concerned, regarding the factors promoting or hindering economic development. It also necessitated identifying discernible trends within the supply side, and ascertaining from those involved in the supply side, both their rationale for involvement and their own perception of the effects their involvement had on economic and social development.

On the demand side tasks needing undertaking included that of identifying the level of consumption of GLAC related goods and services and the trends and explanatory factors behind this level. It would also be necessary to ascertain the extent to which GLAC was used in the business sector, how its use affected the business sector's views of its relevance to economic and social development, and how it effected self confidence in the communities of high Gaelic language use.

Overall it was intended to build up a picture of any long term effect on the economic and social health of the communities brought about through GLAC's impact on intermediate variables linked to regional economic growth, these variables including individuals decisions to stay in, migrate from, or return to these communities; the decisions of incomers to start up businesses, of tourists to visit and externally based firms to locate there.

4.2.1. Defining the Artistic and Cultural Sector

The first task needing to address was an identification of the supply side of GLAC, which would be dependent upon the definition of the sector which was adopted.

In the general field of arts and culture there already existed a useful body of work established by the Policy Studies Institute (PSI) which had undertaken studies of cultural and artistic impacts (albeit mostly in an urban setting).(Rogers 1989; Myerscough 1994; 1995) A study by Bailey and Scott for the Gulbenkian Foundation had also however, considered the question against a rural background, (Bailey and Scott 1989), whilst others had examined the issue of culture from a Gaelic perspective.

Core points of the most useful definitions are outlined in Table 4-1 the implications of which are then discussed below and again by individual component in Table 4-2

Table 4-1 Comparative Approaches to the Artistic and Cultural Sector

General categorisations			**Gaelic categorisations**	
Myerscough	Cultural Trends	Scottish Arts Council	Bryden & N.MacKinnon	K.MacKinnon
Core: *Theatres; concerts; museums; galleries* Ad-hoc companies: *Community organisations and individuals* Cultural Industries: *Sectors supplying artistic products and experiences*	- Authors, composers and 'their own account' artists; - Radio and TV services, theatres etc; - Film production, distribution, and exhibitions; - Manufacture of records and prerecorded tapes; - Printing and publishing of books; - Libraries, museums and art galleries	Music; Drama; Dance and mime; Art; Combined Arts; Literature; Festivals; Touring; Film	Writers and publishers; Artists and galleries; Actors and drama; Musicians and musical performance; Dance; Story telling in oral medium; Local history, craft and non-performing arts	Popular and modern music; Theatre; Cinema; Cultural Festivals; Gaelic Performing Arts; Cultural Policies (local govt); Translation facilities; Gaelic medium education; School materials

Table 4-2 Inclusion of individual components

Component	**Whether included in each model/ approach**				
	Myerscough	Cultural Trends	SAC	Bryden & N.MacKinnon	K.MacKinnon
Theatres;	√	√	√	√	√
Concerts;	√		√	√	√
Museums;	√	√	√	√	√
Galleries	√	√	√	√	√
Community organisations	√		√	√	√
Individuals	√	√	√	√	√
Sectors supplying artistic products and experiences	√	√	√		
Libraries		√	√		
Gaelic Med. Educ.					√
School materials					√
Translation facilities					√
Loc Govt. Cultural policies					√

4.2.2. The Policy Studies Institute Definitions

The PSI studies had adopted two broadly similar working definitions of the artistic/ cultural sector. Myerscough in his publication *The Economic Importance of the Arts in Britain* had defined it as:

> "That part of the economy which has the function of developing and supplying artistic ideas and experiences to the public"
> (Myerscough 1995)

From a similar perspective, and drawing upon Myerscough's earlier works, Rogers in *The Work of Art – a Summary of the Economic Importance of the Arts in Britain* had defined artistic and cultural activities as:

> "Those endeavours in which the development and conveyance of an original artistic idea or experience to the public is the primary purpose, whether to improve or entertain"
> (Rogers 1989)

Myerscough categorised the arts and cultural area into 3 broad sections: a core: (*theatres; concerts; museums; galleries*); secondly, ad-hoc companies: (*community organisations and individuals*); and cultural industries: (*sectors supplying artistic products & experiences*). All the sectors had certain restrictions and exclusions qualifying them.

Also produced by the PSI is *Cultural Trends* (Feist and Hutchison 1989), which concerns itself with 'culture in a fairly narrow sense', including in its remit '*arts, heritage and broadcasting, but excluding fashion and design*.'

Using this approach, they examine trends in 'selected cultural industries': authors, composers and 'their own account' artists; radio and TV services, theatres etc; film production, distribution, and exhibitions; manufacture of records and prerecorded tapes; printing and publishing of books; libraries, museums and art galleries. In addition, they include staff of 'cultural quangos' i.e. The Scottish Arts Council, National Library of Scotland etc., when categorising employment.

4.2.3. The Scottish Arts Council Approach

A third approach which shares much common ground with the studies above may be found in the Scottish Arts Council (1993), in its categorisation for funding purposes. It splits its clientele into the following art-forms: *music; drama; dance and mime; art; combined arts; literature; festivals; touring; film*.

In addition, they also include the 'crafts' economy which occupies an intermediate position between art and design, (although here they acknowledge problems of definition).

Before considering any implications of the above definitions for this thesis, it was also necessary to look at the work done to identify the artistic and cultural sector *specifically associated with Gaelic* in order to investigate whether common ground existed.

4.2.4. Defining the Gaelic artistic/cultural sector

The two most relevant documents in the field using working definitions of the Gaelic arts and cultural sector are - Gaelic Arts towards a New Century (Bryden and

MacKinnon 1993) and Kenneth MacKinnon's report to the Euromosaic Project. (MacKinnon 1994). Macleod (1986) also deals with the issue but in a less systematic way

Brydon and MacKinnon's paper interpret Gaelic Arts as including the following: *writers and publishers; artists and galleries; actors and drama; musicians and musical performance; dance; story telling in oral medium; local history, craft and non-performing arts*

Kenneth MacKinnon in his report to the Euromosaic project categorises the following as included in the Gaelic 'literary production and cultural industry': *popular and modern music; theatre; cinema; cultural festivals; Gaelic performing arts.*

4.2.5. Overlap and points of difference

The overlap and points of difference are as illustrated in Table 4-1 and Table 4-2 above.

As indicated, there is almost complete overlap between the traditional categorisations of the artistic/cultural sector and those seen as pertaining to Gaelic arts and culture.

In *The economics of Gaelic language development* (Sproull and Ashcroft 1993), the approach taken to the supply side of the Gaelic economy *as a whole* is to accept it as:

> "all those activities and jobs whose principle purpose is the provision of Gaelic related goods and services, including the promotion of the Gaelic culture and language"

As can be seen from Table 4-2 this would include the four additional resources *Gaelic Medium Education; School materials; Translation facilities and Local Government Cultural Policies* identified by Kenneth MacKinnon in connection with the reproduction of the *language* itself (i.e. funding of Gaelic medium education etc).

An investigation of the impact of spending on Gaelic medium education was beyond the scope of this thesis however, the thesis focusing on the *artistic and cultural* sector of the Gaelic Economy. In addition, the definitions adopted by the PSI (see above) quite clearly omit the scholastic reproduction of the language itself. Indeed the contemporary categorisation by the Scottish Education Department itself (SED)

treated 'Gaelic Cultural Organisations' under a different category to general 'Gaelic Medium Education' and offered this sector distinctive funding.

The approach adopted in this thesis therefore in relation to supply side considerations has been to construct a model consisting at its core, of a central group of cultural industries/activities, (the centre being based on the common areas outlined in Table 4-1 / Table 4-2). Having adopted these parameters, the research then concentrates on this central group, with the approach to associated provision such as education limited to where a directly cultural aspect could be demonstrably linked to it.

As will be shown below, this 'core' cannot be realistically considered however, without acknowledgment or the role of the funding bodies, whose support framework makes such provision possible. The artistic and cultural support activities of bodies such as Local Enterprise Companies, the Scottish Arts Council, the Gaelic Broadcasting Committee (formerly Gaelic Television Committee), will also be taken into account in investigating the supply side of the picture.

A final point highlighted in later chapters however, (but not obvious from the table above) is the inclusion within the areas studied of the less commercial aspects of Gaelic culture - i.e. story telling in oral medium, the feisean, the amateur festivals etc which the traditional coverage of arts and culture from the PSI and others has tended to ignore. This is of interest in that although its direct financial impact will be seen to be less than more commercial aspects of Gaelic arts and culture (see below, Chapter 5), its overall impact on communities is nevertheless perceived to be substantial.

4.3. Analysing the data

4.3.1. Collecting the data on the supply side

On the supply side, data was obtained from several sources.

Firstly, data was collected from those Gaelic related cultural and artistic establishments and organisations already operating in this sector. During this phase, the decision was also taken to consider the funding bodies which often made such GLAC provision possible - Local Enterprise Companies within areas of the Highlands

and Islands, the Scottish Arts Council and the Gaelic Television Committee. In addition to measuring direct economic impact, this would allow consideration of the views of these bodies when evaluating the economic merit of the 'front-line' GLAC suppliers. Within this area of organisations, the Further Education establishments of Lews Castle College and Sabhal Mor Ostaig were also surveyed, given the unique role of their courses and skill services in relation to GLAC. In this way, it was intended to provide a useful picture of the changing pattern of Gaelic language, arts and cultural development through skills promotion.

The second major aspect of data collection was a survey undertaken of the individual practitioners of art and culture (i.e. ranging from singers, poets, writers, to those involved in the electronic media such as Gaelic Television and Radio) based within the Gaelic economy.

The sampling frame for both these elements of the study was derived from several sources. For organisations the source used was the Arts and Culture sections of the current Directory of Gaelic Organisations, produced by Comunn na Gàidhlig (CNAG)[5], and for individuals, sections of the National Database of individuals involved in Gaelic Arts and Culture held by the National Gaelic Arts Project/ Gaelic Arts Agency (NGAP/GAA)[6], and put at our disposal. Finally, this database was supplemented following discussions held with prominent individuals in the Gaelic artistic and cultural arena who helped fill in omissions where they perceived gaps to lie, based on their personal knowledge of the sector.

4.3.2. Choosing an appropriate method for investigation of the supply side

In considering the research method to be adopted for optimum investigation of this sector the geographically scattered nature of both the organisations and individuals was evident - a fact which would make face to face interviews with the full population impractical in the time scale provided.

[5] Buidhnean Gàidhlig CNAG (1995). Buidhnean Gàidhlig, Comunn na Gaidhlig.
[6]Based in Stornoway, the main organisation dealing with the provision of Gaelic related performers and artistes.

It was decided therefore to utilise a self completion postal questionnaire for both organisations and individuals, which could be followed by a selection of telephone follow ups or face to face interviews where appropriate.

4.3.3. Constructing a questionnaire

Elaboration of the virtues of a well designed questionnaire can be found in most literature on research procedures. It needs to be clear, unambiguous and uniformly workable. Its design must minimise potential errors from respondents, interviewers and coders, and since peoples participation in surveys is voluntary, a questionnaire has to help in engaging their interest, encouraging their co-operation, and eliciting answers as close as possible to the truth.[7]

To this end, careful consideration was given to the format of the questionnaire, including evaluation of the differing attributes of closed and open format questions.

4.3.4. Open and closed questions

May (1995) considers the benefits and disadvantages of open and closed format questions in survey work, and points out that

> "Open questions give respondents a greater freedom to answer the question, because they answer in a way that suits their interpretation[8]...they also have the value of enabling researchers to explore raw data and to devise new coding categories"[9]

Closed questions, on the other hand:

> "limit the possible answers to be given and therefore can be precoded so that each answer may be given a specific number for the purpose of analysis"[10]

In fashioning a questionnaire appropriate to this project, information was to be sought on the following issues:

---

[7]see e.g. Hoinville, G. and R. Jowell (1987). Survey Research Practice. London, Heineman.
[8]May, T. (1995). Social research - issues, methods and process. Buckingham, Open University Press.
[9]SCPR (1981). Survey methods newsletter on Open-Ended Questions, Social and Community Planning Research, May, T. (1995). Social research - issues, methods and process. Buckingham, Open University Press.
[10]May, T. (1995). Social research - issues, methods and process. Buckingham, Open University Press.

The range of supplied services and products

The first area involved the range of services and products supplied by the respondents, and the importance to income/turnover of Gaelic related products within this. The research also sought to identify trends in the pattern of provision of these goods and services, and also trends in the range of customers demanding the services themselves.

Economic activity associated with GLAC

A second area of enquiry concerned the economic activity and employment associated with the Gaelic related aspects of the respondent's activity, including turnover, numbers of employees where applicable and geographical spread of such activity.

Employment issues, dynamic interaction and general opinions

Thirdly, a section on employment issues sought to ascertain the skill mix and qualifications of those employed in this area, including income levels and staff development where applicable. Opinions were then sought on a whole range of dynamic issues, including the rationale behind the respondent's involvement in GLAC, the factors internal and external which aided, or constrained this activity. A last section then asked the respondents views on a battery of issues including opinions of the impact of their and other Gaelic related activity on local culture, self confidence, the possibility of being economically active, levels of tourism, desirability of residence, migration and other questions.

Individual practitioners of GLAC were also asked whether they had any family connections with the Western Isles, or Skye and Lochalsh[11].

Selecting the format of questions

Some of the questions in the supply side survey such as the respondents age range, level of Gaelic, seasonality of employment, ranges of alternative employment opportunities available etc, could be adequately handled by closed questions,

whereas others, dealing with the possibly complicated nature of any dynamic links which might be found between the issues of confidence, feelings of cultural distinctiveness etc, the desirability of maintaining residence in Gaelic speaking areas etc would lend themselves better to an open format treatment, which would allow more flexible coding[12]: Copies of the surveys used in the thesis are found in the appendices .

Piloting the survey

In order to ascertain the suitability of the first draft survey, pilots were sent to a range of individuals and organisations, both those who had as their main focus Gaelic related artistic and cultural work, and others whose work in this medium could be considered as an adjunct to a larger body of non-Gaelic related work.

This resulted in amendment of several areas of the questionnaire, refinement of the open format section, and consideration of the best social and psychological sequence of the questions in order to minimise ambiguity, to better focus the subject area, and to maximise response rates.

4.3.5. Problems inherent in the questionnaire approach

Hughes has pointed out some of the inherent problems caused by the possibility of the researcher and respondent sharing different 'community of meaning structures for assigning cultural significance' to items under research[13]

May has also indicated the differences between what people say they do, and what they actually do (May 1995: p87).

Marsh however (1982: p124) has argued that questionnaires can adequately deal with this issue if a sufficiently complete picture of the context can be obtained through the overall package - additional questions built into the questionnaire can enable the

---

[11] Again this was aimed at better understanding the reasons for which practitioners had become involved in provision of GLAC related goods and to trace whether involvement might be partly linked to family circumstances rather than the recognition of a simple economic opportunity.

[12]"The purpose of coding in surveys is to classify the answers to a question into meaningful categories so as to bring out their essential patterns" Moser, C. A. and G. Kalton (1979). Survey methods in social investigation. London, Gower.

[13]Hughes, J. (1980). The Philosophy of Social Research. London, Longman.

researcher to identify any such variation in researcher/ respondent understanding of the significance of responses. In addition, 'triangulation' - the use of several methods to evaluate the same subject area is also recognised as a valid approach to tackling this problem. These points were borne in mind in the construction of the approach to the issues above.

4.4. Response rates to the supply side questionnaire:

4.4.1. Organisations

Questionnaires were sent out to a wide range of providers of Gaelic artistic and cultural services - 50 in all. Response rates were very high (76%), with the only major omission being Grampian Television. Although information regarding the economic impact of the sector was thus incomplete, the range of response was comparable to previous attempts to gather information from a similar sampling frame (Sproull and Ashcroft 1993). This therefore allowed some confidence in the results.

Table 4-3 Supply side response rates

Organisations:	Returns	Non-returns
Feisean	13	8
Independent TV producers	9	1
The Gaelic book sector	4	0
Gaelic promotion agencies	6	0
Local Enterprise Companies	6[14]	4
Museum/ Arts/ Interp. projects	4	0
The Scottish Arts Council	1	0
Major 'Media'	4	1

[14] Not all LECs were contacted. A small minority were omitted where there had been virtually no Gaelic related tradition in the area covered, i.e. the borders area.

Miscellaneous	3	0

The establishments are delineated in detail in the appendices.

4.4.2. Response of individual practitioners on the supply side

Questionnaires were sent out to almost 300 individuals involved in GLAC activities - the names taken from the NGAP database. A much lower rate of return was received - 53 individuals (17%)

It was recognised from the outset that this would be a more difficult area in terms of response rate. Many of the individuals concerned were involved in this area on a part time basis, or used their activities to supplement income from other sources. Some therefore saw their involvement in GLAC as peripheral, or raising too small a proportion of their income to make the filling in of the complete questionnaire relevant. It may be surmised from comments received from some and also from previous studies that much of this activity is financially unrecorded - being part of the 'underground economy', thus making participants hesitant about responding to detailed requests over income and allied issues.

It was necessary therefore to take into account this low response rate when evaluating the (very) detailed responses which were received and which enable us to add qualitatively to the existing knowledge on attitudes and opinions in this area.

Of the fifty three individuals who replied, a wide range of occupations were reported which are outlined in the appendices.

The majority of the individuals who responded were found however to be based outwith the geographical area of the Gaelic economy[15] (although most conducted some of their economic activity within parts of it). Given the focus on the Gaelic economy within this thesis, normally only the details of GE related activities have

[15] Defined as "the spatial area which stands to gain measurable economic benefits from the further development of the language" (see below p. **Error! Bookmark not defined.**). For reasons outlined in Sproull A, and Ashcroft

been referred to below. Where the use of non-GE details to illustrate a point has taken place, this is made clear in the text.

4.4.3. Collecting the data on the demand side

Two different sources were used to derive data regarding Gaelic related artistic and cultural products and services within the geographical area chosen for the study. These related to individual consumers, and secondly business consumers.

For individual consumers a large scale postal survey was conducted amongst a sample of all individuals within the area concerned.

The second source of data was a detailed interview based survey carried out amongst a sample of businesses in the Western Isles and Skye and Lochalsh

4.4.4. The individual consumer survey

While in an ideal situation, such a survey would have covered the whole population of the area - approximately 36,000 adults, given the availability of Gaelic Radio and TV programmes throughout - such an exercise proved impractical due to its scale and financial implications.

The Electoral Register was therefore used in order to obtain a representative sample of the views of the individuals living in the area.

A postal questionnaire was therefore sent to a randomly chosen sample of 8,400 adults - approximately 25 percent of the adult population. Useable returns were received from 2028 individuals - almost 25 percent of the sample and more than 6 percent of the total adult population of the area. This was a very substantial percentage of the target population with the large response rate possibly being partly explicable through the less financially specific nature of the survey (when compared to the individual supply side practitioner survey)

The ability to infer from a sample to a population as a whole, depends on the scale of the survey and the distribution of key characteristics amongst the population

---

B, The Economics of Gaelic Language Development, 1993, pp 4 – 6, this consists essentially of the Western Isles

surveyed. The scale of the survey and the healthy response rate (25%) suggests that enough data was received to allow statistical inference, provided that no substantial bias was present across key characteristics.

Key characteristics refer to attributes which might positively or negatively bias attitudes to Gaelic related issues in the broadest sense. The literature suggests that variables such as Gaelic fluency, age, gender, and geographical location of residence could be expected to have a bearing on this. The distribution of such variables within the respondents were therefore checked against the most relevant population census for the period in question (Census Report for Western Isles, 1991; Census Report for Highland Region, 1991) and were found to be remarkably close - within 3-4% in most instances. If anything, the percentage difference tended to *under-represent* factors that might be expected to bias the sample positively in favour of the Gaelic language arts and culture. (i.e. see table Table 4-7 Language Competence where those fluent in Gaelic are 'under-represented' by the survey respondents by a factor of 3% in the Western Isles and 9% in Skye and Lochalsh).

---

and parts of Skye and Lochalsh

4.5. Comparison of Survey Respondents and 1991 Census Information

Table 4-4 Geographical Distribution

	Western Isles	**Skye/Lochalsh**
Survey response	69% (1406 respondents)	31% (622)
Actual (over 16s living there)	71.5% (23,330[16])	28.5 (9280[17])

Table 4-5 Gender Distribution

Survey respondents were	45.5% male,	54.5% female
W.Isles population is	49.5 % male,	50.5% female[18]
Skye/Lochalsh pop. is	49% male,	51% female[19]

Table 4-6 Age Distribution

Percentage of survey respondents aged 60 or over	23 %
Percentage of Western Isles population retired (i.e. 60+ for women, 65+ for mer	21%[20]
Percentage of Skye/Lochalsh population retired (as above)	21% [21]

[16]1991 Census Report for the Western Isles, Part 1, p60 Table 8
[17]Based on 100 percent of a 10 percent sample of 928 (1991 Census, report for Highland Region Part 2, P39, Table 72) ;
[18]1991 Census Report for the Western Isles, Part 1, P60, Table 8
[19]1991 Census, report for Highland Region Part 2, P39, Table 72
[20]1991 Census, report for the Western Isles Part 1, Page 21, Table E
[21]1991 Census, report for Highland Region, Part 2, P19, Table E

Table 4-7 Language Competence

Percentage of survey respondents with Gaelic [22](W/I)	69.9%
Actual Percentage of WI population with Gaelic (Age 18+)[2324]	73.3%
Percentage of survey respondents with Gaelic (S/L)[25]	34%
Actual Percentage of S/L population with Gaelic (Age 18+)[26]	43.4%

The content of the individual consumer survey reflected a desire to provide information on 4 questions.

- What was the level and pattern of demand for the full range of Gaelic arts and cultural products and services?
- What factors encouraged or discouraged the consumption of these Gaelic arts and cultural products and services?
- What was the impact of the consumption of each of these products and services on the attitudes and behaviour of the individual consumer?
- What was the general combined impact of consuming Gaelic arts and cultural products on attitudes and perceptions regarding the future of the language and the extent of perceived links between Gaelic and social and economic development?

The collation and analysis of this data would thus allow investigation of the issues identified by the literature review.

The methods used to analyse these results are detailed further below in section 4.6:

[22]Encompassing those who responded that their Gaelic was native/fluent, or 'most everyday conversations'
[23]1991 Census Report for Western Isles Part 1 p 15 plus Table 67, page 139
[24] Latest (2001) Census figures for the Western Isles show an estimated 63.4% of those ages 18+ as having competence in Gaelic. Unfortunately similar figures for Skye/ Lochalsh were as yet unavailable at time of final editing of this thesis
[25]Encompassing those who responded that their Gaelic was native/fluent, or 'most everyday conversations'
[26]1991 Census Report for Highland Region Part 2 p26

4.5.1. The business survey

The other main component of demand for Gaelic arts and cultural products derives from the business sector. Some businesses use Gaelic arts and cultural products in the process of producing or supplying the products or services that constitute their main economic activity. An example of this would be the use of Gaelic music in hotels and restaurants, the use of bi-lingual menus or signage to influence the image of the product etc. While this is likely to be a relatively small component of final total demand for Gaelic arts and cultural products it may be a highly influential one in terms of public attitudes and perceptions about the language-economy linkage.

In addition to identifying the extent to which firms make direct use of Gaelic arts and cultural products in differentiating or improving the overall quality of their products or services, the survey also sought the views of the business community on the overall impact of Gaelic cultural activity on the economic environment within which they operated. Such impacts may operate through a range of channels from altering the image of the area, through attracting more tourists to altering the range of jobs and business opportunities in the area.

To establish the views of the business community on these and other issues a representative sample survey was undertaken of establishments employing 10 or more individuals. These were drawn from all sectors and geographical locations within the study area. In the case of Skye and Lochalsh the sampling frame was supplied by the Local Enterprise Company and in the case of the Western Isles several sources were used, including Digits Young Enterprise Directory, the Western Isles Enterprise Company Directory and the Western Isles Advertising Directory in order to achieve the most comprehensive frame possible in the given circumstances.

37 establishments were contacted in Skye and Lochalsh and 72 in the Western Isles and asked to participate in an extended telephone interview. A response rate of 78% was obtained producing 85 completed interviews. The reasons for non-participation ranged from a flat refusal to difficulties in finding a mutually convenient time to conduct the lengthy interviews, to establishments owned and controlled by a parent company that centrally determined the main substantive issues of interest in the survey.

It was not possible to strictly establish the representativeness of the business survey respondents as regional or sub-regional data on non-micro enterprises (10 or over employees) is incomplete and such data is needed to compare our respondents against. All that can be claimed is that the surveyed firms were chosen at random across all sectors and locations, and that the frame assembled for analysis was the most comprehensive possible in the circumstances.

4.6. Analysing the demand side data (large scale consumer survey)

As a key task of this thesis was the assessment of the relationship between a series of independent variables indicating characteristics of the respondents and a range of dependent variables detailing the respondents consumption of various Gaelic related products and services, a formal method of analysis was required. As will be shown below, the method to be employed was Logit models, with the hypothesis being tested being that consumption of GLAC can be modelled by a number of attributes of the individual.

Formally, this can be written as:

$$C_G = f(Y, L, G, A, S)$$

where C_G = consumption of GLAC related goods and services

Y = income

L = location

G = gender

A = age

S = spoken Gaelic

Testing the Hypothesis

To reach an objective decision as to whether a particular hypothesis is confirmed by a set of data, we must have an objective procedure for either rejecting or accepting that hypothesis.

Siegal (Siegal and Castellan 1988: p7) suggests the following approach to ensure a robust outcome:

- State the null hypothesis H_0 and its alternative (H_1). Decide what data to collect and under what conditions. Choose a statistical test (with its associated statistical model) for testing H_0
- From among the several tests which might be used with a given research design, choose that test, the model of which most closely approximates the conditions of the research in terms of the assumptions on which the test is based.
- Specify a significance level and a sample size
- Find the sampling distribution of the statistical test under the assumption that H_0 is true
- On the basis of (ii), (iii), and (iv) above, define the region of rejection for the statistical test.
- Collect the data. Using the data obtained from the sample(s), compute the value of the test statistic. If that value is in the region of rejection, the decision is to reject H_0; if that value is outside the region of rejection, the decision is that H_0 cannot be rejected at the chosen level of significance.

The results of this approach are outlined below and summarised on Page 147

The null hypothesis and the collection of data

The null hypothesis (H_0) in each of the cases in this section was that possession of the individual characteristics stated has no effect on the respondent's consumption of GLAC goods and services (specifics of each case are stated in the relevant chapter). The alternative (H_1) is that possession of each of these characteristics could be shown to affect consumption. The collection of data and conditions behind the collection have previously been outlined above.

Choosing a suitable statistical test

Amongst the issues to be considered when choosing a method of analysis are:

- Of the methods available, parametric and non-parametric, which uses the information in the sample appropriately - i.e. which test is valid
- Have the assumptions underlying a particular statistical model or test been satisfied
- Are the hypotheses tested by the statistical model appropriate for the situation (Siegal and Castellan 1988:p34)

Statistical tests are often categorised for convenience into parametric and non-parametric depending on the details of the respective populations from which the research sample is drawn.

A parametric statistical test specifies certain conditions about the distribution of responses in the population from which the research sample is taken, whereas a nonparametric statistical test is based on a model that specifies only very general conditions and none regarding the specific form of the distribution from which the sample was drawn.

One of the aspects determining the type of test which is appropriate will be the manner in which the variables are measured. Regression analysis may be an appropriate choice if both independent and dependent variables are ratio or intervally scaled. Discriminant analysis may be chosen when the variables are of a dissimilar nature – the dependent categorical and the independent variables ratio or scaled by interval. Conversely analysis of variance techniques may be used if the independent variables are categorical and the dependent variable is ratio or intervally scaled (Siegal and Castellan 1988).

In the case of this thesis, the nature of the variables used precluded the use of interval, ratio, or even ordinal measures, the independent variables being categorical (i.e. gender; age range etc,) and the dependent variables being dichotomous – i.e. consumed/ did not consume GLAC related goods.

Parametric test were thus inappropriate and it was necessary therefore to examine the non-parametric tests available.

The Chi-Square goodness of fit test

For testing relationships of the type found above when the number of categories are two or more, the chi-square test is often employed. For instance the two way chi-square is appropriate if the subjects, have been cross classified in two ways and the data is in the form of frequency counts(Stevens 1992:p497).

The null hypothesis for a two way chi-square is that the modes of classification are independent i.e.:

H_o = attendance at ceilidhs is independent of respondent's gender and income category

Based on the null hypothesis, expected cell frequencies (e_{ij}) are computed from:

$$e_{ij} = \frac{(rowtotal)(columntotal)}{n}$$ where n = sample size

And compared against the above observed frequencies (0_{ij}), with the following chi square statistic

$$x^2 = \sum \frac{(0_{ij} - e_{ij})^2}{e_{ij}}$$

Although this serves its purpose well with 2 way data, as the data becomes more complex it is increasingly unsatisfactory as a method of investigation as it fails to capture multi-collinear effects and interactions between variables. Consequently this may produce spurious results.

One approach would be to collapse higher order models into less complex ones allowing the analysis of two variables at a time. Although such an approach can be helpful and will provide a certain level of insight into the data, as Dillon and Goldstein point out (Dillon and Goldstein 1984:p316) this also causes a level of compromise i.e.

- the research loses the ability to study interactions involving more than two variables

- the marginal association between 2 variables is quite different from what the association would be in the presence of other variables
- all simultaneous pairwise associations cannot be studied.

Through the use of cross tabs, preliminary analysis of the sample population under study identified the individual characteristics of *Location, Income, Age, Gender* and the *level of Spoken Gaelic* as the main variables that effect the demand for Gaelic live arts and products. In order to formally present these results and to further analyse these findings, cross-tabulations could have been further employed. However this would involve estimation of a considerable number and crucially would not take into account multi-collinear effects as indicated by Dillon and Goldstein above Furthermore, although interactions could be incorporated through controlling for the independent variable, this again would involve producing endless cross-tabs. In order to overcome these two problems, it was therefore decided to examine the use of log-linear (logit) analysis as outlined below, as this would allow all results to be encompassed in a small number of models and importantly provide a framework in which to analyse the data.

Log-Linear Analysis

Log-linear analysis is used when the investigator has several categorical variables (nominal or ordinal variables) and wishes to understand the interrelationships between them. It provides an effective method of analysing categorical data as it does not make any assumptions regarding the independent variables, and is normally recommended when the independent variables do not satisfy the 'multivariate normality' assumption. (Sharma 1996:p317)

Log-linear analysis helps in understanding patterns of associations among categorical variables. It is applicable to two-way tables but is more commonly used for three-way or larger tables where interpretation is more difficult. Although there is no theoretical limit on the number of categorical variables which can be analysed, in practice three or four is most common, with rarely more than 5 analysed concurrently (Afifi and Clark 1998:p410). Experience within the current study has also identified sample size as a limiting factor, and that above 4 or 5 variables the usefulness of the framework in analysing the data is considerably reduced.

Analysing two-way tables and above - notation and procedure

In order to use chi-square to test a single sample, it is necessary to calculate the expected frequencies, in order to test the null hypothesis of independence.

The expected frequency is:

$$\mu_{ij} = N\pi_{ij}$$

where μ and N denote population parameters, i denotes the ith row and j the jth column.

Theoretically, if two events are independent then the probability of them occurring is the product of their individual probabilities.

In the case of a two-way table, the expected frequency in the ijth cell when H_0 of independence is true is given by:

$$\mu_{ij} = N\pi_{i+}\pi_{+j}$$

(where $i+$ represents the sum of the ith row, and $+j$ represents the sum of the jth column, therefore π_{i+} and π_{+j} represent the marginal probability in the ith row and jth column respectively)

From chi-square to two way log-linear analysis

The log-linear model is, as its name implies, a linear model of the natural logs of the population parameters.

Proceeding from the previous equation, and taking the log of μ_{ij} in the case of a confirmed null hypothesis, we then have:

$$\log\mu_{ij} = \log N + \log\pi_{i+} + \log\pi_{+j}$$

since the log of a product of three quantities is the sum of the logs of the three quantities.

This linear model is usually written as:

$Log\lambda_{ij} = \mu + \lambda_i^A + \lambda_j^B$ (independence model)

Where A denotes the row variables and B its column variables and the λ's are used to denote either Logμ or Logπ

Given that there are many possible values of the λ's, in order to make them unique two constraints are commonly used, namely that the sum of the$\lambda_i^A = 0$ and the sum of the $\lambda_j^B = 0$

If the null hypothesis is not true, then it is necessary to adjust the log-linear model.

To show an association between the row and column variables a fourth parameter is added to the model to denote this:

$Log\lambda_{ij} = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$ (association model)

where λ_{ij}^{AB} measures the degree of association between the rows and columns.

Note that the association terms are also required to add up to zero across rows and across columns. This latter model is sometimes called the **saturated** model because it contains as many free parameters as there are cells in the interior of the frequency table.

Testing the fit of the model

Two goodness of fit statistics are available to test the model - the Pearson chi-square goodness of fit test and the likelihood ratio chi-square test.

The Pearson chi-square test is the same as the regular chi-square test of independence namely:

$$x^2 = \sum \frac{(observedfrequency - expectedfrequency)^2}{expectedfrequency}$$

where the observed frequency is the actual frequency and the expected frequency is computed under the null hypothesis of independence from the marginals in the table.

This is done for each cell and then summed for all cells. Within the log-linear model, the model's predictions of expected frequencies are used to run the goodness of fit.

To perform the test of no association, the expected frequencies are first computed without the inclusion of the λ_{ij}^{AB} term in the model (independence model above).

If a small chi-square value is obtained this illustrates a small difference between expected and observed values and hence indicates a good fit. There is therefore no need to include the term measuring association - the independence model being sufficient to explain the data.

The aim is to find the simplest model that will yield a reasonably good fit. If the model fits, the aim then is to find an even simpler model (parsimonious) that is almost as good. What is sought after is the simplest model which fits

In general if the model does not fit (i.e. the chi-square is too large), more terms are added to the model.

Another test which it is possible to use is the likelihood ratio chi-square (G^2) which can be written as:

$$G^2 = 2\sum (observedfrequency)\log\frac{(observedfrequency)}{(\exp ectedfrequency)}$$

Again the larger the value of G^2, the poorer the fit, and the weaker the null hypothesis.

Although similar to the Pearson chi-square, G^2 has the added advantage that it is additive in hierarchical models such as those adopted in the examples examined for this research thus indicating the relative difference between models.

Although Afifi (Afifi and Clark 1998:p418) has noted that Pearson chi-square may give better results if some of the cells have small frequencies, the size of the sample allowed the use of the likelihood ratio chi-square.

In order to interpret the size of chi-square, it is necessary to know the degrees of freedom associated with each of the computed chi-squares. The degrees of freedom for the goodness of fit chi-square are given by the number of cells in the interior of the table minus the number of free parameters being estimated in the model. This can be very complicated to calculate in higher than two-way tables, but fortunately is done automatically in the SPSS programme used for analysis.

Three-way and higher models

The research undertaken for this thesis consists of a series of two-way and three way log-linear models (see below).

Construction of higher level log-linear models proceeds in a similar manner to that for a two-way as indicated below (although with an increasing degree of complexity). This can be seen in the structure of a three-way model for instance, several examples of which are now outlined:

Complete independence model

In this case all variables are unassociated with each other, and in a three-way model, this is shown by:

$$Log\square_{ijk} = \square + \square_i^A + \square_j^B + \square_k^C$$

This model includes no parameters that signify any association

One variable independent of the other two

If a third variable C is jointly independent of both variable A and variable B, then the only variables that can be associated are A and B. Hence only parameters denoting the association of A and B need be included in the model. This allows us to write

$$Log\square_{ijk} = \square + \square_i^A + \square_j^B + \square_k^C + \square_{ij}^{AB}$$

Note that similarly, any of the variables could be jointly independent of the second and third.

Conditional independence model

This refers to a model where A and B are independent of each other across all layers (*k* layers) of C. but where associations between variables A and C and variables B and C, exist.

This model can be written as:

$$Log\lambda_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$$

Note that similarly other conditional independence models can be produced with the omission of the A and C association term or the B and C association term.

Models with all two-way associations

It is also possible to have a model with all three two-way associations present. This is called the homogenous association model (Afifi and Clark 1998: p413)and would be written as:

$$Log\lambda_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$$

In the above case all three two-way associations are included.

Saturated model

Finally the saturated model, where all associations are included may be illustrated by:

$$Log\lambda_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$$

Choosing the model

Given a high number of possible models to examine, it is common for researchers to restrict their analysis to that of hierarchical models, which are found to be easier to interpret than non-hierarchical ones given that the results of each effect are known to be consistent across all levels, thus allowing accurate comparisons (Afifi and Clark 1998: p424). Hierarchical models adopt the rule that whenever the model contains higher order interactions with certain variables, it also contains the lower order

effects with these same variables, i.e. if λ_{ij}^{AB} is included in the model, then λ, λ_i^A and λ_j^B must be included. As Afifi has indicated, this helps restrict the models to those which are easier to explain. This therefore has been the approach adopted in this research.

In order to restrict the number of variables and interactions within hierarchical models which it is necessary to examine, two methods can be employed - forward selection or backwards elimination. Backwards elimination (which features in the SPSS programme for log-linear modelling) is normally used when a single sample (cross sectional survey) is being examined and there is no prior knowledge of how the variables may be interrelated.

Procedurally, the chi-square statistics are examined, starting with high order interaction terms included, which will usually have a small value of chi-square, indicating an acceptable model. The aim then is to eliminate non-significant items, and still find a model which has a good fit. If less complex models can be found which do retain a good fit then the aim has been achieved. Although within this thesis, as many as six of the interactive models were found to be saturated (compared to eight unsaturated, or parsimonious models), this procedure remains successful as the saturated models indicate all the variables which are of importance and the direction in which they operate.

The relationship between log-linear and logit models

Log-linear models treat all the variables equally and attempt to clarify the interrelationships between them. However, in many situations the investigator may consider one of the variables as an outcome or dependent and the rest of the variables as explanatory or independent. In the case of this thesis, the hypothesis under test was that consumption of GLAC can be modelled by a number of attributes of the individual.

Logit models are a class of log-linear models[27] used to investigate a relationship between a dichotomous dependent variable (such as consume/ not consume) and independent categorical variables – i.e. gender; age range etc,. Like log-linear analysis it is a comparatively new approach compared to other techniques of analysis and thus there is not such a body of models as exists with regression analysis (Afifi and Clark 1998:p441). Nevertheless, useful examples do already exist of the appropriate use of this technique e.g. (Jobber, Mirza et al. 1991; Jobber 1999)

Unlike the case in discriminant analysis where calculations are carried out on a dependent variable coded as having a value of 0 or 1, in a logit model, the value of the dependent variable is based upon the 'log odds' . In this case, the independent variable is used to explain the variation of this measure.

The 'odds ratio' is the ratio between the frequency of being in one category and the frequency of not being in that category.

For example, if 500 people in a sample attend a Gaelic ceilidh and 250 do not, then the odds of a person attending are 500/250 = 2.00, or 2 to 1.

For a two by two table, the odds ratio is estimated as the product of the observed diagonal frequencies, divided by the product of the non diagonal frequencies. i.e.:

$$\theta = \frac{(O_{11}O_{22})}{(O_{12}O_{21})}$$

if $\theta = 1$, then the variables are independent.

In logit analysis, the log is taken of this ratio (the 'logit') and analysis then takes place on this measure and the impact of the independent variables.

In the case where the outcome variable is dichotomous (e.g. consumes/ does not consume Gaelic related products), we can denote the outcome variable by y

[27] In general, it can be shown that any logit model is equivalent to a logistic regression model. Afifi, A. A. and V. Clark (1998). Computer-Aided Multivariate Analysis. London, Chapman & Hall.

Once we know the probability for one value y_1, the probability of the other value (y_2) is determined, because the sum of the probabilities must be 1

i.e. if we denote the probability of the first value by $P(y_1)$ and the probability for the 2nd value by $P(y_2)$, then the odds of y_1 occurring are given by:

$$\text{odds} = \frac{P(y_1)}{1 - P(y_1)} = \frac{P(y_1)}{P(y_2)}$$

In practice it is found more convenient to work with the log of the odds ratio, known as the 'logit'.

If we denote the logit by Ω, then

$$\Omega = \ln(\text{odds}) = \ln \frac{P(y_1)}{1 - P(y_1)}$$

In using the logit to describe the relationship amongst cells, if the logits in some cells are found to be roughly equal, then this means that the conditions characterised by these cells are not having a differential effect on the outcome variable.

Where the logits for some cells are substantially different, the conditions characterised by the cells *are* having a differential effect on the outcome variable.

Limitations of Logit Models

Even Logit models however have limitations. Most importantly is that there is a limit to the number of variables that can be included in a single model for a given sample size. This is because the addition of each variable effectively results in cutting the sample in half for a dichotomy, a third for a three category variable and so on. For example, if the model contained a dichotomous dependent variable and four trichotomous variables then this would effectively divide the sample into 192 individual cells. If the sample size was under 2000, then in several of the class very few observations would remain in which to estimate a parameter. This can lead to a high number of insignificant parameters being included in the model (importantly due to a high standard error rather than an insignificant parameter estimate per se) and also spurious and unreliable results.

In order to overcome this limitation, the approach adopted here is to estimate two models for each Gaelic art event, a general model and an interactive model. The former includes those variables that have been identified as having few, if any, interactions, hence their effect on consumption will be estimated and presented. The interactive model includes variables with strong interactions, and again their effect will be estimated and the results analysed

4.7. A note on interpreting the tables in Chapter 6

Each model is normally presented in the form of two tables – firstly a table consisting of the General and Interactive Logit models (see above) and secondly a table of odds ratios (if two or three way interactions have been found to be present). As mentioned above, the main effects being tested are location, income, gender, spoken Gaelic and age. Given the dichotomous nature of the first four, the following notation has been used:

Location(1) = 'Urban/town' (as opposed to 'rural')

Income(1) = 'Below median income' (as opposed to 'above median income')

Gender(1) = 'Male' (as opposed to 'female')

Spoken Gaelic(1) = 'Can speak Gaelic' as opposed to 'cannot speak any Gaelic'

In terms of Age categories,

Age(1)= 'young' (aged 18 – 35);

Age(2)= 'middle aged' (aged 36 – 59)

Age(3)[28] = 'older' (aged 60+)

Reading the General and Interactive Models tables

In the tables representing the General and Interactive Logit Models , the odds of the main effects (i.e. the factors given above), are shown in the last column. These are the

odds of consumption/attendance as compared to the sample average[29] .For example, Table 6-3, page 200, Main Effect Location (1), (someone living in an urban/town location) has odds of 0.84 of attendance compared to the sample average.

The odds of attending compared to someone living in a rural setting (i.e. the dichotomy of urban), would of course be different[30] – and where appropriate this latter measure is used from time to time for illustration purposes in the analysis.

The column of Z values are the standard significance levels (i.e. 1.96 at 5%), thus allowing a judgement as to whether results should be accepted or rejected. For the purposes of the models examined, a significance level of 10% (0.1), has normally been regarded as satisfactory if the fit of the model is thus improved. The 'Estimate' column refers to the parameter estimate given by SPSS, the exponential of which provides the figure within the odds column.

Interpreting the Odds Ratios tables

When two way (or three way) interactions take place in the General and Interactive Model, their odds need further manipulation before use to show how the main effects work over the different layers of the model. Thus, rather than using the odds stated in the two way interactions and three way interactions rows in the General and Interactive Models, attention instead should be given to the 'Odds ratios' tables, which illustrate how each effect works over each layers of the interaction.

Within the Odds Ratios tables, the odds referred to are not against the sample averages, but against the relevant opposite dichotomous variable.

Thus in Table 6-4, p. 202, the two way Gender/ Age odds ratio, tells us that amongst the younger category (18 – 35) of attendees, males are more likely to attend than females by a factor of 1.20 , in the middle aged category they are less likely to attend to a factor of 0.72 etc[31]

---

[28] This group is not shown in the table, but may be calculated from Age 1 & 2 odds.
[29] And are calculated by the exponential of the parameter estimate produced by SPSS
[30] Calculated by the square of the odds against the sample average
[31] These odds ratios are found by calculating the exponential of the sum of the relevant parameter estimates, for the two characteristics under examination and then dividing one by the other to get whichever odds are sought.

4.8. Testing the Hypothesis - summary

Null Hypothesis H_0 = Characteristics of respondents such as gender, income, age etc have no effect on consumption of Gaelic related goods and services. H_1 = Such characteristics have an impact on consumption of such goods.

Statistical Test. Log-linear (logit) analysis is chosen given the categorical nature of the independent variables and the dichotomous nature of the categorical dependent variables.

Significance level. Let $\alpha = 0.10$, and let N = 2008, the number of replies to the questionnaire

Sampling distribution . Under the null hypothesis, x^2 as computed from $x^2 = \sum \frac{(0_{ij} - e_{ij})^2}{e_{ij}}$ is distributed approximately as x^2 with degrees of freedom $df = (r-1)(k-1)$, where r = the number of rows and k = the number of columns in the contingency table. When H_o is true, the probability associated with the occurrence of values as large as, or larger than an observed x^2 is shown in a table of critical values of the chi-square distribution

Rejection region. The region of rejection consists of all values of x^2 which are so large that the probability associated with their occurrence when H_0 is true is less than or equal to $\alpha = 0.10$

The detailed outcome of this approach is found in **Chapter 6** below.

For example, to find the odds ratio of the two way association of Young men and Young women, it is necessary to calculate firstly
Exponential of the sum of (Intercept parameter estimate + Male parameter estimate + Young parameter estimate + Young/Male parameter estimate).
This is then divided by the Exponential of the sum of (Intercept parameter estimate + Female parameter estimate + Young parameter estimate + Young/Female parameter estimate). A similar procedure is applied to the other two way interactions to calculate the odds ratios. The procedure for three way interactions is of a similar nature, but adding the parameters for the three way associations.