

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**A MONTE CARLO STUDY OF
THE EFFECT OF SAMPLE BIAS ON THE MULTINOMIAL
LOGIT CHOICE MODEL COEFFICIENTS.**

**A Thesis presented in partial fulfilment
of the requirements of the
degree of Master of Business Studies at Massey University.**

Grant K. Bell

1996

ABSTRACT

This thesis reports the findings of a Monte Carlo simulation into the effect of sample bias on the parameters of the multinomial logit (MNL) choice model. At issue is the generalisability of parameter estimates obtained from biased samples to the balance of the population. An actual data set of 164 respondents was used to estimate an aggregate model. Using these parameters as the true coefficients of choice behaviour, an unbiased sampling distribution of the MNL parameters was derived by repeatedly fitting aggregate models to artificially generated sets of individual responses. Subsequently, the biased sampling distribution was derived by selectively eliminating those individuals at the tails of the sample distribution based on their correlation with one of the independent variables.

The expected values of the biased and unbiased sampling distributions were compared to assess the sensitivity of the model to sample bias. The research found the biased coefficients changed by significantly more than the proportion of individuals removed. However, this sensitivity was predictable as the percentage change in the value of the coefficients was related to the size of the coefficient. It was also found that the coefficients of the unbiased variables were not significantly influenced by bias on another variable. The ratio between the unbiased variables was also maintained. It was concluded that although sensitive to bias, the estimates produced by the MNL model could be modified to reflect the different effect of the bias on the coefficients. Additionally, there was no evidence to suggest that the MNL estimates were not reflecting the effects of interest when calibrated on unbiased samples.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisors, Michael Fox and Malcolm Wright, for their insight, professionalism, and patience.

The support, financial and otherwise, provided by the Marketing Department was also greatly appreciated.

Special thanks to my family whose support made this possible.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	ii
TABLE OF CONTENTS	iii
LIST OF TABLES AND FIGURES	vi
1.0 INTRODUCTION	1
2.0 LITERATURE REVIEW	3
2.1 The Multinomial Logit Choice Model	5
2.2 Treatment of Errors	7
2.3 The Scale Parameter	10
2.4 MNL Estimation	11
2.5 Data Sources: Stated and Revealed Preference	14
2.6 Empirical Applications of the MNL	17
2.7 Sources of Error in the Estimation of the MNL Choice Model	20
2.7.1 Specification Error	21
Omitted Relevant Variable	21
Included Irrelevant Variable	24
Cross-sectional Variation in Preferences	25
2.7.2 Breaches of the MNL's Assumptions	30
Independence from Irrelevant Alternatives	30
Independently and Identically Distributed Errors	32

2.7.3	Measurement Error	34
2.7.4	Aggregation Bias	35
2.7.5	Systematic Sampling Error (or Bias)	37
2.7.6	Random Sampling Error	38
2.7.7	Summary	40
2.8	Monte Carlo Research	41
3.0	OBJECTIVES	44
4.0	METHOD	47
4.1	Experimental Design and Calibrated MNL Model	50
4.2	Generating Individuals and Samples	54
4.3	Biassing the Sample	56
4.4	Generating the Unbiased Sampling Distribution	57
4.5	Generating the Biased Sampling Distribution	59
4.6	FORTRAN 77 Computer Programme	62
4.7	Scale Parameter Normalisation	63
5.0	RESULTS AND DISCUSSION	
5.1	Recovery of the Original Parameter Estimates	67
5.2	Impact of Sample Bias Upon the Parameter Estimates of the Biased Variables	70
5.3	Impact of Sample Bias Upon the Parameter Estimates of the Unbiased Variables	77
5.4	Sensitivity of the Unbiased Variables to Sampling Bias	80
5.5	Theoretical Standard Errors versus Standard Deviation of the Unbiased Sampling Distribution	82
6.0	CONCLUSION	85

7.0 REFERENCES 89

8.0 APPENDICES

Appendix A: FORTRAN 77 programme 103

LIST OF TABLES AND FIGURES

Table		Page
1.	Design of the 12 Choice Sets in the 2 ⁹ Factorial Main Effects Vacation Destination Choice Experiment	51
2.	MNL Model of Vacation Choice Behaviour Estimated on the Original Set of 164 Individuals	52
3.	How well does the Monte Carlo method recover the original coefficients?	68
4.	How much does bias affect the value of the biased variable?	71
5.	Does the affect of the bias depend upon the value of the coefficient of the biased variable?	72
6.	Nonlinear regression: Is the size of the coefficient a good predictor of the effect of sample bias?	75
7.	How much does bias on one variable affect the value other variables?	78
8.	How sensitive is each variable to bias in other variables?	81
9.	Are the Asymptotic Standard Errors Accurate?	83
Figure		
1.	Plot of Coefficient with Percentage Effect of Bias	73

1.0 INTRODUCTION

The multinomial logit (MNL) model is used by marketers to predict consumer product or brand choice behaviour as a function of that brand's or product's attributes and the characteristics of the consumer. The MNL is a random utility model which assumes that choice consists of both a systematic (or explainable) component and a random component. It is widely used in industry, to guide strategy, and in academia, where it furnishes results which test specific knowledge claims about consumer behaviour.

However, a number of potential sources of error exist in the estimation of the coefficients of the model. These include specification error, breaches of the assumptions underlying the model (namely, independence from irrelevant alternatives, IIA, and independently and identically distributed errors, IID), measurement error, aggregation bias, random sampling error, and systematic sampling error (or bias). The first five of these have already been investigated by a number of authors including Gordon, Lin, Osberg, and Phipps (1994), Ben-Akiva and Lerman (1989), Jones and Landwehr (1988), Batsell and Polking (1985), Lee (1982), Horowitz (1981), and Chamberlain (1980). Systematic sampling error or bias has received the least attention. This is surprising given the fact that the MNL is nearly always calibrated on samples drawn from the population of interest.

In marketing, three of the more common sources of data used to estimate the MNL include retail scanner data, consumer panels, and experimental data. This data either describes actual market behaviour with associated product attributes and consumer characteristics or hypothetical choice behaviour with the alternative attributes and actual consumer characteristics of a selected sample of individuals.

The main repercussion of calibrating the MNL on samples is the requirement that at the very least the model functions as expected in conditions of sampling error. It would appear that research

specifically directed at examining the stability or otherwise of the MNL model in situations where sample bias is prevalent has not been attempted. This study investigates the behaviour of the MNL coefficients when estimated on biased samples. In particular, the Monte Carlo method is employed to generate biased sampling distributions which are compared with a benchmark unbiased sampling distribution. The bias is simulated by removing individuals who are most (or least) highly correlated with one particular independent variable.

Three main consequences of this simulated sample bias are explored. The effect on the biased variable is examined to determine if the coefficients vary by the same proportion as the bias. Ideally, the MNL's biased coefficients would change by a proportionately less amount. This would mean that the MNL is producing estimates that reflect the effect being modelled more than the sampling error. If the coefficients change by proportionately more than the bias, then the stability of the model would be questioned.

Another effect of interest is the change in the unbiased variables caused by the simulated bias. It would be desirable (and expected) for the unbiased variables to remain unchanged. As the Monte Carlo method used here generates individuals with no interactions between the independent variables, we would not expect any significant change to occur. However, if this does occur, then it would be a major handicap to the MNL.

Finally, any change in the coefficients of the unbiased variables should not significantly impact on the ratio between them. If this does occur, then the difference in the coefficient sizes are both a reflection of the effect and the error. The MNL would therefore be unjustifiably sensitive to bias.

If the MNL is found to be sensitive to sample bias, then not only is the collection of an unbiased sample important, but the assumptions underlying the model may be dubious.

2.0 LITERATURE REVIEW

In marketing, discrete choice models (also known as quantal, qualitative, or categorical models) are used to predict future demand for a product or service. These models' predictions are consistent with consumers employing a specific set of decision rules in aggregate when purchasing products or services and are made by measuring the impact of attributes on consumer choice between alternative products or services. Discrete choice models were initially developed by psychologists, such as Thurstone (1927), whose comparative judgment models of individual and group behaviour initiated the concept of random or stochastic utility (Amemiya, 1981, p 1490; McFadden, 1986, p 279). They have since been employed in a wide variety of disciplines from medicine to accounting.

Two broad classes of discrete choice models exist. Compensatory models estimate consumer demand by assuming that consumers trade-off product or service attributes : "...good features of an alternative compensate for the bad" (Johnson, Meyer and Ghose, 1989, p 255) . In contrast, noncompensatory models do not make this assumption and include decision rules where consumer choice depends on the product possessing one important attribute or the fact that a product does not contain one unsatisfactory attribute. Compensatory models have dominated the marketing literature with the multinomial logit model being the most used. This is in spite of evidence that consumer decisions are to some degree based on noncompensatory strategies (Johnson, Meyer and Ghose, 1989, p 255). The main reason given for the use of compensatory models is their ability to mimick a wide range of noncompensatory strategies although the research of Johnson, Meyer and Ghose (1989) suggests that in negatively correlated (non-orthogonal) research designs, this is not always the case.

Choice models can be estimated at three levels: aggregate; disaggregate; and individual. An aggregate model is one estimated on macro level data. In marketing, this could be at segment level, such as geographic location or income level, or at the level of the total market (as

represented by the aggregated responses of the individuals in the sample). Disaggregate models are estimated on micro level data, with the model being fitted to the individuals' responses to all the choice sets rather than their aggregated responses. Individual level models are fitted to separate individual's responses.

Aggregate models are generally used to gain predictions of total market response and are preferable to disaggregate models as they provide a better fit to the data and the estimates contain less error. Furthermore, providing a sensible base for aggregation of the sample is selected (i.e. one that restricts heterogeneity), aggregate models allow for the estimation of more effects or levels for the same budget and sample size as a similar disaggregate model. However, when the sample population is heterogeneous with respect to their choices across segments, then aggregation may be dubious and a disaggregate model would provide more accurate estimates of effects. The main reason for estimating individual level models is explanation of individual consumer behaviour.

2.1 The Multinomial Logit Choice Model

The multinomial logit (MNL) model was derived from a set of axioms about the choice process first developed by the mathematical psychologist Luce (1959). The model is fundamentally a regression model designed to treat discrete data. A discrete variable is one which takes only discrete values 0, 1, 2, ..., n and not fractional quantities. This type of data is particularly relevant to the marketing environment where products, in the form of bundles of attributes, are offered to consumers. The model calculates the probability that an individual or a group of individuals will select an alternative as a function of the attributes of the alternatives in the choice set.

The MNL model relies on several assumptions. The first assumption is that each alternative j in the total set of available alternatives S_n has utility U_j for the individual k , equal to a deterministic component V_j and a random component ϵ_j :

$$U_{jk} = V_{jk} + \epsilon_{jk}.$$

The deterministic component is calculated as a function of the observed attributes of the alternatives and/or the observed characteristics of the respondent. This represents the effect of these observables on the choice process. For the MNL, the deterministic component V_{jk} equals:

$$V_{jk} = \sum_{i=1}^I \beta_{ijk} X_{ijk}, \quad \text{where: } i \text{ is the } i\text{th attribute of alternative } j, \\ \beta \text{ is the coefficient of attribute } i, \text{ and} \\ X \text{ is the level of attribute } i \text{ on alternative } j.$$

The random component represents the unobserved variables that impact upon choice from one occasion to the next as well as the stochastic error inherent in the choice process.

Secondly, it is presumed that the individual chooses the alternative with the highest utility from the set of available alternatives on that occasion. Thirdly, the probability of any of the utilities being equal is zero: $U_j \neq U_n$.

Therefore, the probability of individual k choosing j equals:

$$\begin{aligned} P_k(j) &= \Pr \{U_j \geq U_n, n \in S_n\} \\ &= \Pr \{V_j + \varepsilon_j \geq V_n + \varepsilon_n, n \in S_n\}. \end{aligned}$$

Finally, the random components or errors are assumed to be independently and identically Gumbel distributed:

$$\Pr(\varepsilon_{jk} \leq \varepsilon_{nk}) = e^{-e^{-\varepsilon}}, \quad -\infty < \varepsilon < \infty.$$

Together, the four assumptions above combine to produce the MNL model, where $P_n(j)$ is the probability that individual (or aggregated sample) n chooses alternative j , μ is an arbitrary scale parameter, and V is the deterministic component of the utility function associated with the observable characteristics or attributes of the decision maker or alternative:

$$P_k(j) = \frac{e^{\mu V_{jk}}}{\sum_{n \in S_n} e^{\mu V_{nk}}}.$$

In most cases, due to its computational attractiveness, the utility function is assumed to be linear-in-parameters (Ben-Akiva and Lerman, 1989, p 108). A linear-parameters utility function equals:

$$U_{jk} = \sum \beta_{1jk} X_{1jk} + \beta_{2jk} X_{2jk} + \beta_{3jk} X_{3jk} + \dots + \beta_{ijk} X_{ijk} + \varepsilon_{jk};$$

although a multiplicative utility function is sometimes used:

$$U_{jk} = \sum \beta_{1jk} X_{1jk} \times \beta_{2jk} X_{2jk} \times \beta_{3jk} X_{3jk} \times \dots \times \beta_{ijk} X_{ijk} \times \varepsilon_{jk};$$

Nevertheless, this condition is not as restrictive as it would seem as the coefficients can be transformed to improve the fit of the model by better representing the relationship between the dependent and independent variables (Amemiya, 1981).

2.2 Treatment of Errors

The main distinguishing feature between the MNL and other choice models such as the linear probability (LP) model and the multinomial probit (MNP) model is the way in which the errors (ϵ) are treated.

The LP model assumes an independent and identical “uniform” distribution of the random components ($\epsilon_j - \epsilon_n$). The model is usually estimated through ordinary least squares (OLS). Estimates obtained through OLS are consistent and unbiased but not asymptotically efficient due to heteroscedasticity in the variances of the errors. This heteroscedasticity can be removed and an efficient estimator attained by dividing the independent variables and the constant by the standard deviation of the error variance (Hensher and Johnson, 1981, p 166). The resulting estimator, weighted least squares (WLS) is consistent, unbiased, and asymptotically efficient.

Nevertheless, the OLS and WLS estimates of the LP model do not always produce predictions that rest between zero and one as they theoretically should. Furthermore, the OLS and WLS estimators of the LP model suffer from their inability to yield normally distributed errors. The consequence of this is that the coefficients cannot be subjected to statistical tests (Hensher and Johnson, 1981, p 168).

The MNP model relaxes the assumption of independence from irrelevant alternatives (discussed later in this section) and imposes a normal distribution on the disturbances (ϵ_j and ϵ_n). It achieves this by allowing each individual's response to the product attributes to vary about a mean value. The MNP also permits correlation between the errors of different alternatives. This can be seen by examining the utility expression for the MNP (adapted from Hensher and Johnson (1981, p 184):

$$U_{jk} = \sum_{i=1}^I \beta_{ij} X_{ijk} + \left(\sum_{i=1}^I \gamma_{ik} X_{ijk} + \epsilon_{jk} \right),$$

where:	i	=	the i th attribute of alternative j
	j	=	alternative j in choice set S_n
	k	=	the k th individual
	U_{jk}	=	the utility of alternative j for individual k
	β_{ij}	=	the mean effect of attribute i on alternative j
	X_{ijk}	=	the attribute description i of alternative j for individual k
	γ_{ik}	=	the individual taste variation of the k th individual to the i th attribute of alternative k
	ϵ_{jk}	=	the stochastic error in k th individuals response to alternative j .

Each individual's response to the variable X_{ijk} is allowed to vary by γ_{ik} around the mean β_{ij} representing individual taste variation. This results from the correlation between the errors of different alternatives. Also, the structure of the error term assumes the ϵ_{jk} are normally distributed enabling covariance between alternatives.

The MNL model is preferred to the MNP for computation reasons, the normal distribution requiring estimates of a location parameter, μ , and a variance parameter, σ , whereas the MNL requires only the former as the variance is fixed. In practice, this is significant for any problems requiring the estimates of more than two or three attributes (i) and alternatives (j). For example, a MNL problem where $i = 4$ and $j = 4$ would only require four estimates of the $i \beta$'s. However, the MNP would also need $i(i+1)/2$ (ten) estimates for the γ_{ik} , and $j(j+1)/2$ (ten) estimates of the elements of the covariance matrix for the ϵ_{jk} (Hensher and Johnson, 1981, p 185).

The MNL was developed to provide a tractable model for situations where estimation of the MNP would be infeasible. In order to achieve this the MNL model assumes Gumbel distributed errors, which does not allow for correlation between the errors of different alternatives or for covariance in the systematic components of the alternatives. The direct result of this restriction is the independence from irrelevant alternatives (IIA) property. The MNL model computes the

probability of choice for a particular alternative by determining the ratio of that alternative's utility to the sum of all utilities. Therefore, the addition of a new alternative to a choice situation should not impact on the ratio between the alternatives prior to the new alternatives inclusion. As Hensher and Johnson (1981) state, "this property is valid only if the new alternative competes *equally* with each existing alternative" (p 149). The IIA assumption is further discussed in section 2.6.2.

The Gumbel distribution of the MNL errors is also called the Weibull, the double exponential, or the type II extreme value distribution and is a close approximation of the normal distribution but with fatter tails. A number of authors claim that estimates produced by the MNL model do not differ markedly from those produced by the multinomial probit except for large samples (Chambers and Cox, 1967; Amemiya, 1981, p 1487). The Monte Carlo study of Gordon, Lin, Osberg, and Phipps (1994) supported this to some extent. They found the probit estimates to be approximately 2.5 percent higher than logit estimates for sample sizes of less than 4,000 observations. However, beyond this the difference between the probit and logit estimates increased to about four percent for sample sizes up to 20,000.

The normal and Gumbel distributions, of the MNP and MNL models respectively, are considered to be behaviourally superior to the uniform distribution of the LP model. The LP model produces probability predictions that are linear in the interval between the systematic components of the alternatives ($V_j - V_n$). Outside of this interval, the LP model predicts either a zero or 100 probability of choice. Consequently, the LP model has the problem that it produces predictions that contradict the sample observations. This primarily involves alternatives where the model predicts a zero probability of selection, when in actuality the sample contains observations which select that alternative (Ben-Akiva and Lerman, 1989, p 94). The MNP and MNL models avoid this problem by imposing normal and Gumbel distributions respectively on the errors. The repercussion of this is the more behaviourally acceptable condition that the predicted probabilities outside of interval between the estimated systematic components ($V_j - V_n$) tend asymptotically toward zero and one.

2.3 The Scale Parameter

The scale parameter has an important role in the value of predictions made by the MNL and is an inherent characteristic of random utility models. In the MNL, the random component is Gumbel distributed. The Gumbel distribution has a fixed variance equal to $\pi^2/6\mu^2$, where μ is the scale parameter (Ben-Akiva and Lerman, 1989, p 104).

This feature of the Gumbel distribution impacts upon the probability estimates made by the MNL as the size of the parameter estimates dictates the amount of variance in the utility function explained by the systematic component. As the value of the scale increases, the β 's are transformed into progressively larger values, and vice versa. Therefore, large values of the μ (and consequently the β 's) lead to the systematic component dominating the random component; small values of the μ (and therefore the β 's) see the random component dominating the systematic component (Chen and Anderson, 1993, p 238-39).

The effect of this on the choice probabilities is marked. When the systematic component (V) dominates the random component (ϵ), the utilities show greater levels of dispersion. As a result, the difference between the higher and lower attributes as measured by the β 's increases and therefore, the MNL choice probabilities "...predict deterministically to the alternative with the highest systematic utility" (Swait and Louviere, 1993, p 306). When the random component dominates the systematic, the choice probabilities approach the uniform distribution, so each alternative will have the same probability of being selected (Chapman and Staelin, 1982).

This problem is compounded by the fact that the scale parameter cannot be calculated for any one model due to "...confounding with the vector of utility parameters. Hence, it is arbitrarily set to one when dealing with any given model" (Swait and Louviere, 1993, p 305). In spite of this, in the case where the same vector of parameters are estimated on different data sets, the value of the ratio of the scale parameters can be determined between the models (Swait and Louviere, 1993).

2.4 MNL Estimation

The parameters $\beta_1, \beta_2, \beta_3, \dots, \beta_n$ of the logit model can be calculated from a sample of observations through the use of a number of estimators: maximum likelihood; minimum chi square, which is also called weighted least squares, Berkson, or Berkson-Theil; minimum Pearson chi square; maximum score; nonlinear least squares; and ordinary least squares. In marketing, the maximum likelihood (ML) estimator has been used in most applications though the minimum chi square has also been applied (Bunch and Batsell, 1989).

Maximum likelihood estimates are those set of parameters which would produce the sample of observations more frequently than other possible estimates (Hensher and Johnson, 1981, p 43). The estimates are gained in an iterative fashion by maximising the log of the likelihood function, \mathcal{L}^* :

$$\mathcal{L}(\beta_1, \beta_2, \dots, \beta_i) = \sum_{k=1}^K \sum_{j \in S_n} y_{jk} (\beta' x_{jk} - \ln \sum_{n \in S_n} e^{\beta' x_{nk}})$$

- where:
- β_i = the estimated coefficient of attribute i
 - k = the kth individual in the total sample K
 - j = the jth alternative
 - S_n = the set of alternatives
 - y_{jk} = the individual k's response to alternative j
 - $\beta' x_{jk}$ = the vector of estimates of the attributes of alternative j
 - n = the nth alternative
 - $\beta' x_{nk}$ = the vector of estimates of the attributes of alternative n

(Ben-Akiva and Lerman, 1989, p 118).

The log likelihood function is globally concave so if a solution exists, it will be unique. This is in

contrast to some estimators where a unique solution is not assured. Maximum likelihood estimates are consistent, asymptotically normal, and asymptotically efficient (Ben-Akiva and Lerman, 1989, p 22). A consistent estimator is one whose variance around the true value of the population parameter decreases as the sample size increases. An asymptotically normal estimator is one where the distribution of the estimates approaches the normal distribution as the sample size tends to ∞ . An asymptotically efficient estimator is one whose variance “is less than or equal to that of any other estimator” (Ben-Akiva and Lerman, 1989, p 29).

However, Ben-Akiva and Lerman (1989) point out that these are large sample properties and “maximum likelihood estimators are not in general unbiased or efficient” (p 22). A number of authors (McFadden, 1974, p 123; Malhotra, 1984, p 22 ;Ben-Akiva and Lerman, 1989, p 22; Chen and Anderson, 1993, p 238) also suggest that the small sample properties of the ML estimator are not well known. Malhotra (1984) recommends that for small samples ($n < 50$) the ML estimator may be “inappropriate” (p 22). Unfortunately, Malhotra does not offer any empirical or statistical reasons for his definition of small samples.

Chen and Anderson’s (1993) Monte Carlo study looked specifically at the small sample properties of the ML estimator with sample sizes ranging from 20 to 200. They found that the precision of the estimates improved as the degrees of freedom increased. Consistent with the study of Horowitz (1981) they also found this improvement occurred at a decreasing rate. Furthermore, Chen and Anderson (1993) observed that increases in the number of choice alternatives and smaller scale parameter values improved the precision of the estimates. This study apart, little empirical research exists to define the small sample properties of MNL estimators. However, this should not be detrimental to the findings of this study if the sampling distributions produced are large.

The minimum chi square (MCS) estimator involves setting the value of the β ’s such that the relationship between the dependent variable, the log odds ($\ln (P_k(j)/P_k(n))$), and the independent product attributes is maximised. It is asymptotically equivalent to ML but generally not as

computationally tractable as MCS estimators require a large number of observations per cell (Ben-Akiva and Lerman, 1989, p 95-97). Additionally, if cells equal zero or if all respondents respond positively, an arbitrary modification of the cell values is required (Malhotra, 1984, p 21). Again for small samples, the MCS estimator may be unsuitable as the small sample properties of MCS are not known (Malhotra, 1984, p 22).

The minimum Pearson chi square (MPC) is asymptotically equivalent to ML and MCS but is not as computationally attractive as the β 's have to be determined through nonlinear optimization (Bunch and Batsell, 1989, p 59). The maximum score estimator (Manski, 1986) is generally not as efficient as other estimators, though in circumstances where the disturbances are heteroscedastic, it is preferable to other estimators such as ML and MCS which assume homoscedasticity. Nonlinear least squares (NLS) and ordinary least squares (OLS) are consistent but not asymptotically ML which would be preferable (Bunch and Batsell, 1989, p 59). The NLS estimator also requires nonlinear optimization to be solved.

In the most comprehensive study of estimators in a marketing environment, Bunch and Batsell (1989) found ML to be superior to MCS, MPC, NLS, and OLS. They based this conclusion on a comparison of the vector of Monte Carlo estimates produced by the different estimators with the true parameter vector. They also compared the ability of each of the estimators to predict the true probabilities in a holdout sample, and gauged the reliability of the statistical inferences made by the different estimators through an examination of the asymptotic normality of the estimates. Their results contradicted the findings of a number of previous studies (Berkson, 1955; Amemiya, 1980) which suggested that MCS was superior to ML. However, they corroborated the study of Smith, Salvin, and Robertson (1984) with regard to the validity of the statistical inferences of ML. Bunch and Batsell (1989) reasoned that previous studies which found that MCS was superior to ML could be attributed to the bioassay nature of the experimental conditions which contrast to marketing applications which consist of more alternatives in the choice set and greater numbers of explanatory variables. This finding provides tentative support for the use of ML in this study.

2.5 Data Sources: Stated and Revealed Preference

For a MNL model to provide estimates of the trade-off between attributes, data needs to be collected that captures not only the choice made by an individual or population of interest, but also the attributes of the available alternatives and/or the characteristics of the consumers who selected that alternative. Two separate streams of data have been used to achieve this: revealed preference (RP) data which describes actual consumer behaviour (Ben-Akiva, Bradley, Morikawa, Benjamin, Novak, Oppewal, and Rao, 1994) and stated preference (SP) data which imitates choice by exposing respondents to hypothetical choice sets. Models based on SP data therefore reflect behavioural intentions as they don't require respondents to change their actual behaviour (Adamowicz, Louviere, and Williams, 1994).

Two of the more prevalent forms of RP data include retail scanner data and consumer panels. Scanner data can be collected at the individual, store, or aggregate level. For individual data, each purchase made of a product is assumed to have been made by a separate individual from the set of available products in the category. These individuals can then be aggregated. Store and aggregate level scanner data represents the total sales of each brand in a product category. Consumer panels take many forms but generally yield information on the purchase behaviour of individuals or households, their characteristics (e.g. number of individuals in the household), and the nature of the purchase environment (e.g. competing brands, relative price). Information on purchase behaviour can either be collected directly by the researcher from scanner information or indirectly from the respondent. The individual or household information from the consumer panel can be aggregated for models of total market response.

Although RP data is restricted to "observations on past or present actual market choice behaviour" (Ben-Akiva et al, 1994), SP data takes many more forms including:

Stated Intention - expected future choice behaviour

Stated Choice	-	hypothetical choice behaviour
Stated Judgment Ratings	-	overall attractiveness of actual or hypothetical choice alternatives considered one at a time
Attribute Ratings	-	attractiveness of various attributes of real or hypothetical choice alternatives
Similarity Ratings	-	perceived amount of similarity between pairs of real or hypothetical choice alternatives either overall or attribute by attribute
Attitudinal Ratings	-	respondents feelings about a market situation

(Ben-Akiva et al, 1994, p 336).

Despite the range of methods for collecting SP data, caution should be taken. Ben-Akiva et al (1994) suggest that the context and format of the choice sets has an affect on the response (p 337). Furthermore, they state: "Artificially framed SP tasks, such as rating, ranking, or trade-off exercises, tend to detract from the validity of survey responses" (Ben-Akiva et al, 1994, p 337). The results of McFadden and Leonard (1992) would support this claim as they found the model's estimates to be sensitive to different formats, questions and information provided (Ben-Akiva et al, 1994, p 337).

The validity of SP data also presents a problem that is not inherent in RP data. SP methods are criticised because actual behaviour is not observed or may not change (Adamowicz et al, 1994, p 337). Therefore, models fitted to SP data may in fact be producing predictions of preferences or intentions rather than actual behaviour.

However, RP data suffers from a few major disadvantages, mainly (1) the need to make an assumption about the structure of preferences which in some cases can not be tested, (2) colinearity in the independent variables, and (3) a limited range of attributes and in particular, levels (Adamowicz et al, 1994; Ben-Akiva et al, 1994).

Consequently, SP methods are used to overcome all these problems. Furthermore, combining both SP and RP data has become popular and presents an opportunity to develop more powerful models. Ben-Akiva et al (1994) and Louviere and Swait (1993) discuss the issues related to combining multiple data sources while Adamowicz et al (1994) present an empirical example of this. The approach of Adamowicz et al (1994) involves comparing the MNL predicted probabilities of the RP and SP parameter estimates to ascertain if they originate from the same preference structure. They found a distinct linear relationship between the predictions of the two data sets and could not reject the null hypothesis that the parameters are equal. A log likelihood ratio test (see Swait and Louviere, 1993) was then performed to determine if the scale parameters of the two data sets were different. They found that the scale parameters were different at the 99 percent level of confidence. The RP and SP data sets were then concatenated and the SP data rescaled relative to the RP data. This was achieved by finding a constant value that maximised the joint log likelihood of the combined data sets but that did not produce a log likelihood for the combined model that was significantly different from the separate log likelihoods of the RP and SP MNL models. The joint model parameter estimates could then be used.

2.6 Empirical Applications of the MNL

The fact that the MNL can treat discrete data and is computationally attractive compared to the LP and MNP models has meant the logit model has been widely used for a number of different applications in a variety of disciplines. These include:

Medicine:

where Marshall, Shroyer, and Grover (1994), Agresti (1993), and Chester, Kaufman, and Wolff (1992) looked at issues relating to the impact of different treatments on patient recovery and risk assessment in coronary bypass operations.

Education:

in which Kerkvliet (1994), Dey and Astin (1993), Leppel (1993), Bunn, Caudill, and Gropper (1992), Kodde (1986), Fuller, Manski, and Wise (1982), and Bishop (1977) investigated problems relating to college enrolment decisions, cheating behaviour, and education demand.

Agriculture/Horticulture:

where Linden, Rita, and Svojala (1996), Garcia, Woodard, and Lee (1995), Kanninen and Khawaja (1995), Kling and Herriges (1995), Ferguson and Yee (1995), and Bell, Roberts, and Park (1994) analysed questions relating to salmon fishing, lethal temperatures in apple, prediction of human caused forest fires, participation in ag/hort programmes, and statistical issues relevant to ag/hort.

Real Estate:

Iwarere and Williams (1991), Hayashi, Isobe, and Tomita (1986), and Jud (1983) looked at the issues of industry location behaviour, residential housing demand, and tenure choice.

Finance/Banking:

where Lawrence and Arshadi (1995), Aguilera and Gonzalez (1993), Billensley and Thompson (1985), Ederington (1985), Hasbrouck (1985), and Marsh (1982) examined loan repayment, problem loan resolution, bond ratings, bank stock repurchases, debt/equity choice of companies, and the characteristics of takeover targets.

Geography:

where Krieg (1993), Khraif (1992), Nijkam and Reggiani (1992), and Waddell (1992) investigated problems of migration, and race and urban structure.

Human Resource Management:

Lewis (1986), Borjas (1984), Kolchin and Hyclak (1984), Berger, Olsen, and Boudreau (1983), Killingsworth and Reimers (1983), Vanderporten and Hall (1983), Ehrenberg and Marcus (1982), Mauro (1982), Perloff (1981), and Gunderson (1980) analysed the problems of labour force participation, gender and promotion, labour turnover, union participation, effect of unions on job satisfaction, race differences and pay/promotion, collective pay negotiations, minimum wages, strikes, and unemployment.

Transportation:

Mandel, Gaudry, and Rothengatter (1994), Glaister (1983), Lioukas (1982), Levin (1978), and Thompson (1977) confronted issues relating to rail commuter demand, transportation planning, freight transportation, and transit route structure.

Economics:

where economists such as Jung (1993), Tzeng, Chen, and Lan (1991), Asch, Malkiel, and Quandt (1986), Cecchetti (1986), Dalton and Marcis (1986), Lo (1986), Morgan (1986), Cameron (1985), Fields and Mitchell (1984), Grossman (1984), Stewart, Harris, and Carleton (1984), Hannan (1983), Shields and Tsui (1983), Banerjee (1982), Falaris (1982), Finn and Rappoport (1982), Greenspan and Vogel (1982), Hughes and

McCormick (1981), and Figlewski (1979) have found the MNL model particularly amenable to the problems they face. These have included analysis of energy conservation and demand, job satisfaction, pricing, corporate bankruptcies, the impact of government policy, merger behaviour, migration effects, market entry decisions, birth probabilities, imports, debt indicators, and market efficiency.

And Marketing:

where Gensch and Soofi (1995), Horowitz and Louviere (1995), Neelankawil, Mummalaneni, and Sessions (1995), Chintagunta (1994), Gupta and Chintagunta (1994), Jobber (1994), Fader (1992), Gonul and Srinvasen (1993), Horowitz and Louviere (1993), Perkins and Roundy (1993), Swait and Louviere (1993), Fader, Lattin, and Little (1992), Johnson, Meyer, and Ghose (1989), Jones and Landwehr (1988), Gatignon and Reibstein (1986), Malhotra (1984), Guadagni and Little (1983), Louviere and Hensher (1983), Louviere and Woodworth (1983), Chapman and Staelin (1982), Currim (1981), Gensch and Recker (1979), Green, Carmone, and Wachpress (1977) and many others have investigated a number of issues including consideration sets, advertising copy, brand positioning, estimation and estimators, segmentation, heterogeneity, tests of competing specifications, experimental design, the use of different sources of data, the scale parameter, and forecasting of demand for different products and brands in different markets.

2.7 Sources of Error in the Estimation of the MNL Choice Model

Error in the process of estimating the coefficients of discrete choice models arises in varying degrees from six sources: specification error; breaches of model assumptions; measurement error; aggregation bias; random sampling error; and systematic sampling error or bias (adapted from Hensher and Johnson, 1981, p 224-226). Each of the six sources “can be difficult to separate out in an analysis of errors” (Hensher and Johnson, 1981, p 224), especially, for example, breaches of the assumptions implicit in the model, some of which can be solved through the addition of explanatory variables, and thus could be considered as specification error. These errors impact not only on the estimated coefficients but also the standard errors associated with these and the ensuing forecasts in the form of choice probabilities.

2.7.1 Specification Error

When constructing a model at the individual or aggregate level, one cannot expect to capture the behavioural process in its entirety. Theoretic, economic, and practical considerations impact on the degree to which a model can be precisely specified. For example, household interactions account for variation in the choice of breakfast cereal, but due to a lack of theoretical or empirical information, may not be able to be included in a model. Such simplification of the processes influencing actual behaviour are reflected in specification error (Hensher and Johnson, 1981, p 224).

Five main types of misspecification may be present in any MNL model: omitted structure where relevant explanatory variables are excluded (Lee, 1982; Horowitz, 1981; Manski, 1973); inclusion of irrelevant variables (Horowitz, 1981); cross-sectional variation in preferences where the functional form of the utility varies across the sample which includes heterogeneity bias (Jones and Landwehr, 1988; Horowitz, 1981; Chamberlain, 1980; Manski, 1973); instrumental variables which are unsuitable surrogates of actual variables (Hensher and Johnson, 1981, p 224; Manski, 1973), and; poorly defined alternatives (Hensher and Johnson, 1981, p 225). This review discusses the first three in detail as they seem to have been the most researched.

Omitted Relevant Variables

Analysis of omitted structure has concentrated on the omission of relevant variables and the inclusion of irrelevant variables. This is of particular importance to this study as the sample bias simulated in this study is the same as partially omitting a variable. In his mathematical analysis of omitted variable bias, Lee (1982) investigated the conditions upon which the omission of a relevant explanatory variable will not produce bias in the estimates of remaining variables. Lee compared a correctly specified two variable model with a misspecified one variable model which had one of the original variables removed. Two propositions were forwarded and subsequently

supported by a series of proofs.

The first proposition suggested that bias would not occur if the omitted variable z was independent of the included variable x , conditional on the response variable y (Lee, 1982, p 199). This is similar to the linear regression model except that the independence between the omitted variable z and the included variable x is not conditional on the dependent variable (Lee, 1982, p 208). However, Lee (1982) does indicate in a footnote that the constant term will be biased but does not investigate this further (p 203).

The second proposition explored the direction of bias when the “omitted variable z is dichotomous, and the included variable x is discrete” (Lee, 1982, p 207). In the linear regression model, the direction of bias depends upon the sign of the coefficient of the omitted variable and the correlation of the included and omitted variable (Lee, 1982, p 208). For the logit model, Lee (1982) proved that the coefficient of the included variable x would be:

- “ (i) biased upward if either $\beta_i > 0$ and $\delta_k > 0$, or $\beta_i < 0$ and $\delta_k < 0$,
(ii) biased downward if either $\beta_i > 0$ and $\delta_k < 0$, $\beta_i < 0$ and $\delta_k > 0$,
(iii) unaffected if either $\beta_i = 0$ or $\delta_k = 0$ ” (p 207)”,

where δ_k is a measure of the association between the omitted and included variable conditional on the dependent variable.

Horowitz (1981) examined the impact of the omission of a relevant explanatory variable on the choice probabilities produced by the logit model. He concluded that inconsistent forecasts would result from the exclusion of a relevant variable unless one of five conditions is satisfied:

- “ (1) The omitted variable must be distributed independently of the included variable.
(2) Either the omitted variable must have equal mean values in all alternatives, or else alternative specific constants must be included in the utility function to represent

the effects of the alternative specific means of the omitted variable.

- (3) Either the omitted variable must be IID across alternatives or else a more general model must be used.
- (4) The omitted variable must have the same distribution in the population for which the forecasts are made as in the population from which the values of the models parameters are estimated.
- (5) The omission of the relevant variable must not substantially alter the parametric form of the random component of the utility function.” (Horowitz, 1981, p 427).

Furthermore, Horowitz (1981) used Monte Carlo simulation to investigate this problem. He assumed that the correct model was $V(X_i, Y_i) = 5X_i + Y_i$ with the misspecified model being $V(X_i) = aX_i$. With an absolute standard deviation of 0.32 and a maximum absolute error of 0.85, he concluded that the logit model was particularly sensitive to omission of relevant variables (p 434). This is surprising as the stochastic error term is supposed to capture some of this specification error, and the estimation of the effects was independent. Nevertheless, Horowitz (1981) may not have controlled for the scale parameter differences which could explain this surprising result.

Another form of specification error studied by Horowitz (1981) was “correlated random utility components and explanatory variables” (p 428). This is a unique case of omitted variable bias where its occurrence results in inconsistent parameters and choice probabilities (Horowitz, 1981, p 428). For example, consider a model of cereal choice, where package size is an independent variable with the utility function equal to $U_{jk} = a_{ijk}X_{ijk} + \epsilon_{jk}$. If family size (an omitted variable) impacts on choice and is correlated with package size, then ϵ_{jk} and X_{ijk} , will also be correlated. If family size was included in the model then the ϵ_{jk} and X_{ijk} would not be correlated.

The only available methods at present to control for this effect are to discover the causes of the correlation and account for these variables in the specification or to use a more general model such as probit. Despite this, the logit model appears to be robust against correlated utility components. In his Monte Carlo example, Horowitz (1981) compared the true probit model

where the utility function $U(X_i, Y_i) = 5X_i + Y_i + \varepsilon_i$ and 0.95 correlation between ε_1 and ε_2 with a logit model where the utility function $V(X_i, Y_i) = aX_i + bY_i$. The root mean square difference between the probability forecasts produced by the logit model and the true probit model equaled 0.022 with a maximum error of 0.074. Nevertheless, this result should be treated with caution as the scale parameter was not accounted for.

Included Irrelevant Variables

The impact of including irrelevant variables was also discussed by Horowitz. An irrelevant explanatory variable is one with a utility coefficient of zero. Horowitz (1981) suggested that the maximum likelihood estimation procedure produces consistent estimates of the utility parameters and as a result will not cause inconsistent or biased estimates (p 421). Additionally, though the parameters will not be affected, sampling errors will increase as a result (Horowitz, 1981, p 421). However, increases in the value of the sampling errors is restricted to models with linear-in-parameters utility functions.

For nonlinear-in-parameters models, inconsistent estimates can be produced but this “depends on whether there exist values of the parameters of the erroneous model that cause the true and erroneous models to coincide” (Horowitz, 1981, p 421). If these exist, sampling errors will increase but the model will remain consistent. If not, the model will produce inconsistent forecasts. For example, if the true model of choice equals $U = \alpha X$, an incorrectly specified model $U = aX + bY$ will yield consistent estimates whereas the incorrectly specified model $U = aX + bY^2$ will yield inconsistent estimates (Horowitz, 1981, p 428). Horowitz (1981) presents an example of the increase in the sampling error as a result of including an irrelevant variable (pp 421-22). He compared the sampling error in a true model where $P_1 = 1 / (1 + \exp \alpha)$ with an irrelevant variable model where $P_1 = 1 / [1 + \exp (a + b X)]$ and the irrelevant variable is X . The resulting sampling errors in the misspecified model were twice that of the true model.

Cross-sectional Variation in Preferences

There has been comparatively little analysis of the problems relevant to cross-sectional variation in preferences in the MNL model. This more than likely represents acceptance that solutions to this problem are statistically cumbersome. Heterogeneity bias has received most attention with studies of particular relevance being those by Chamberlain (1980), Horowitz (1981), and Jones and Landwehr (1988). Jones and Landwehr (1988) state that “not accounting for heterogeneity when estimating logit choice models...may lead to biased parameter estimates and more severely biased choice probability estimates” (p 41).

Chamberlain (1980) was one of the first to control for heterogeneity in homogenous models although the logit model was not particularly considered. He developed three algorithms for analysing grouped data: the joint likelihood function, the conditional likelihood function, and the marginal likelihood function. Chamberlain was interested in estimating the parameters common to all groups. But the effect of incidental parameters in fixed effects homogeneous models such as the basic logit model are not estimated and therefore, biased estimates can result (Chamberlain, 1980, p 225). The MNP, by allowing correlation between alternatives, avoids this problem.

Chamberlain (1980) attempts to develop an estimation procedure that estimates incidental parameters (α_i) which capture the group effects and removes bias from the fixed parameter estimates common to all groups. The primary objective was to find estimators of β that converged as the number of groups increased, even if the observations in each group were small (Chamberlain, 1980, p 236).

Chamberlain found that the joint likelihood maximum likelihood estimator produced inconsistent parameter estimates. Estimates produced by the conditional likelihood function were consistent. The conditional likelihood function exploits the independence from irrelevant alternatives assumption by conditioning on sufficient statistics for the incidental parameters (i.e. the intercept). In other words, the MNL estimates are conditional on the shared or similar intercept of the

individuals.

The main advantage of this approach is that the likelihood function does not rely on the incidental parameters and so standard asymptotic theory for maximum likelihood estimation applies. The final method investigated was the marginal likelihood function which proved to produce consistent estimates. With this approach, the incidental parameters are assumed to follow a distribution. As discussed below, this approach can however result in more biased estimates than those produced by the joint likelihood function (Jones and Landwehr, 1988, p 43).

Horowitz (1981) referred to heterogeneity bias as “random taste variation” (p 424). In the MNP utility function $U_{jk} = \beta_{ij} X_{ijk} + (\mu_{ik} X_{ijk} + \varepsilon_{jk})$, the individual taste parameters are represented by μ_{ik} . If these fluctuate randomly in a manner that is not accounted for by the X_i 's, then the parameter estimates will also vary randomly (Horowitz, 1981, p 424). Consequently, the random taste variations will cause the logit model to produce inconsistent estimates of parameters and choice probabilities (Horowitz, 1981, p 425). However, Horowitz (1979) found that in the case of linear-in-parameters utility functions, the logit yields consistent estimates of the mean values of the taste parameters but inconsistent choice probabilities. Horowitz (1981) suggests the only methods available to account for this form of error are to add more variables to the utility function to account for individual taste variation or to use a more general model such as probit (p 426).

In his simulation study, Horowitz (1981) compared a three alternative probit model with the utility function specified as $U(X_i, Y_i) = \alpha X_i + \beta Y_i + \varepsilon_i$ with an incorrectly specified logit model $V(X_i, Y_i) = aX_i + bY_i$ (p 432). As mentioned in section 2.1, the logit model does not allow for random taste variation. With a maximum absolute error of 0.18 between the predicted probabilities of the logit and probit models, Horowitz (1981) judged the choice probabilities estimated by the logit model to be sensitive to random taste variations. Again, it would appear that Horowitz (1981) did not control for the scale parameter.

Jones and Landwehr (1988) extended the research of Chamberlain by removing heterogeneity bias

from logit model estimation through the use of a conditional likelihood function. They contend that Morrison (1966) first implemented the concept of heterogeneity by proposing to model it with a probability distribution. This was achieved by modelling the distribution of the household specific intercept terms (Jones and Landwehr, 1988, p 42). In contrast, Jones and Landwehr estimate specific values for the household intercepts. They do this based on the research of Heckman and Singer (1982) which suggested that structural models are extremely sensitive to assumed choices of the distribution of the unobservables. Jones and Landwehr (1988) do not impose any distribution on the household specific intercepts and thus avoid biasing the estimates of the parameters (p 43).

There are three methods for treating heterogeneity in the logit model (Jones and Landwehr, 1988, p 42):

- | | | |
|-------------------------------------|---|---|
| Homogeneity | - | both household specific intercepts and parameters are constant. |
| Explainable Heterogeneity | - | both household specific intercepts and parameters are constant with variation in choice across households being explained by variables used to describe households. |
| Household-Specific
Heterogeneity | - | intercept varies across households while parameters remain constant. |

Jones and Landwehr adopt the household-specific heterogeneity approach to remove bias. They contend that “there is variation in preference between households that cannot be explained by knowing sufficient amounts of household specific data” (p 42). Their paper uses Chamberlain’s (1980) conditional logit estimation technique and extends it to variable length purchase strings and multinomial choice. Within the context of a two parameter model, Jones and Landwehr compare the results produced by the traditional homogenous maximum likelihood estimation with those of the heterogeneous conditional maximum likelihood technique.

They compared the fit as measured by the proportion of explained variance of the two techniques and found the conditional method with a fit of 0.81 (n=200) was superior to the homogenous model (fit = 0.11, n=3205). The parameter estimates for the two methods were also compared. For the Last Purchase coefficient, both models predicted relatively similar values of 2.045 for the conditional approach and 2.008 for the homogenous method. In both cases this parameter was significant at the 0.01 level. The Relative Price coefficients exhibited greater variation. The conditional logit estimate of 0.380 was less than that returned by the homogenous model (0.641). Additionally, a predictive test of the models was undertaken. They compared the two models' predictions and the actual shares of Brand A for each week of the data and found little difference between the homogeneous and heterogeneous models.

Jones and Landwehr (1988) concluded that the heterogeneous model is most appropriate when explanation, not prediction is the primary purpose of the model (p 55). This is due to the household-specific intercept terms which are only estimated for the sample and cannot be transferred to the population. The fixed effects model is more suitable when estimating the impact of changes in the explanatory variables on the probability of purchase (Jones and Landwehr, 1988, p 55).

However, the heterogeneous model assumes if a household does not purchase a brand in the estimation set, then it never purchases that brand. Additionally, if a household always purchases the same brand, then it is always assumed to purchase that brand. Households like this could be removed from the data set. However, there are two consequences of removing households that never switch brands from the data set (Jones and Landwehr, 1988, p 55). The first is that the model is based on evoked sets rather than all possible brands. Secondly, the variables that affect a consumers choice on the margin such as relative price and promotions should be included. But variables that affect whether the brand will be purchased at all (e.g. product quality, household income) should not be included as the model is predicting if the brand will be purchased on the next occasion due to these variables rather than whether it will be purchased at all. Clearly, the method that can be used to account for heterogeneity depends upon the purpose and type of

model being constructed.

2.7.2 Breaches of the MNL's assumptions

Arguably, the two main assumptions underlying the Luce-based random utility models such as the MNL model are independence from irrelevant alternatives (IIA) and independently and identically distributed (IID) random utility components. Violation of either of these impacts upon the precision of the estimates and forecasts produced.

Independence from Irrelevant Alternatives (IIA)

The most scrutinized assumption of the logit model is the independence from irrelevant alternatives axiom. This is a direct corollary of the conditional probability axioms developed by Luce (1959). In essence, IIA means that if alternatives are added to or removed from a choice set, then the ratios of the choice probabilities between alternatives should remain the same. It is explained best by Ben-Akiva and Lerman (1989), "for any two alternatives, the ratio of their choice probabilities is independent of the systematic utility of any other alternatives in the choice set" (p 130). This assumption is advantageous in terms of computational tractability over more general models such as probit or elimination-by-aspects. Of course, IIA can be considerably unrealistic.

The regularly cited blue bus/red bus problem is an example. Suppose that a commuter is faced with the choice of either a car or a blue bus to travel to work. The consumer assigns a probability of 0.5 to each. If an additional choice of a red bus was added to the set, then it would be conceivable that the consumer would still travel by car 50 percent of the time and 25 percent each on blue bus or red bus. However, IIA would be violated in this case and a Luce type model would assign values of one-third to each. It is therefore important that alternatives presented in the choice sets are distinct (McFadden, 1974, p 113; Ben-Akiva and Lerman, 1989, p 52) though what is meant by distinct can only be determined by comparing the estimates from the full data set and subsets.

Batsell and Polking (1985) have identified three main problems with the IIA assumption: the similarity problem; the dominance problem; and violation of regularity (p 178). The similarity problem was illustrated above in the red bus/blue bus example. The dominance problem occurs when one product dominates another. Batsell and Polking (1985) give the example of a market containing three products, Coke, Pepsi and 7-Up. For all choices between a cola and a 7-Up, the market shares are equal 0.5 and 0.5. However, when confronted with a choice between the two colas, Coke is dominant with 95 percent share. A Luce-type model cannot capture this dominance. Violation of regularity arises when the choice ratios are not allowed to change. In a widely used example, a choice between two bottles of wine, a consumer may be indifferent. But if the shop owner adds another bottle which is the same as the first bottle, the consumer may favour the single bottle over one of the identical pair. A Luce model is constrained to maintain the constant ratio between the alternatives. Of course, this problem could be viewed as misspecification due to failure to include a "uniqueness" variable.

Most academic attention in this area has focused on providing researchers with statistical tests to ascertain the legitimacy of the IIA assumption for the sampled choices. McFadden, Tye, and Train (1977) explored the issue of IIA tests and this research was subsequently extended by Small and Hsiao (1982) and Hausman and McFadden (1984). Tests of IIA involve comparing the parameter estimates from the full data set with those from models estimated on subsets of alternatives (Ben-Akiva and Lerman, 1989, p 184). If IIA is a legitimate assumption, then the model estimated on the full set should be similar to that gained from restricted sets. The most effective of these tests are discussed by Ben-Akiva and Lerman (1989) and all involve the comparison of the log likelihood value adjusted parameter estimates from the full sample with those from the restricted sets.

The IIA problem can be avoided through the use of the nested logit model. The nested logit can be used when the alternatives share some unobserved and observed attributes. For example, for a model of biscuit choice, the dependent variable choice is predicted as a function of package size, price, and stochastic error. However, the choice alternatives include both chocolate and non-

chocolate brands. If a joint model was estimated, then IIA would be breached. A nested logit model allows for covariance between the utilities of subsets of alternatives with a common attribute (Ben-Akiva and Lerman, 1989, p 286). This is achieved by dividing the total error into two independent components, shared (package size and price) and non-shared (chocolate and non-chocolate). The model is estimated sequentially.

Independently and Identically Distributed (IID) Errors

In the multinomial logit model, the structure of the random utility components is assumed to be independently and identically distributed across the alternatives. The assumption that the errors are independently and identically Gumbel distributed places a key restraint on the values of the utilities. The fact that the random components are Gumbel distributed is not considered to be a major restriction as this approximates the normal distribution of the MNP model (Ben-Akiva and Lerman, 1989, p 104) though some empirical evidence suggests that different values of the parameters are estimated (Gordon et al, 1994). Nevertheless, IID does constrain the ensuing errors to have the same scale parameter which leads to the assumption of “homoscedastic disturbances” (Ben-Akiva and Lerman, 1989, p 107). The repercussion of this is that each disturbance has the same variance which is clearly not always the case. Ben-Akiva and Lerman (1989) discuss tests of heteroscedastic random utilities which are a violation of the IID assumption of homoscedastic disturbances.

For example, consumer choice of breakfast cereal may be influenced to some extent by brand considerations. If consumers prefer Kellogg’s cereals, then a model that did not account for this factor in its specification would violate IID. The random utility of non-Kellogg’s cereals would tend to be negatively correlated with each other and the Kellogg’s cereals positively correlated. The mistaken assumption that the random utility components are IID leads to an incorrectly specified relationship between the independent variables and the choice probabilities and consequently, biased estimates and forecasts (Horowitz, 1981, p 423).

In spite of this, the assumption of independence is not entirely groundless. "This assumption may not always be met in reality since the unobservable ϵ 's may be correlated among certain individuals. However, this assumption is made...for the sake of convenience, since it would be usually impossible to specify a correct correlation structure" (Amemiya, 1981, p 1491).

Horowitz (1981) addressed the problem of non-IID random utility components and found the logit model to be stable in their presence (p 434). The correct three alternative probit model with utility function $U(X_i, Y_i) = 5X_i + Y_i + \epsilon_i$ was compared with a logit model $V(X_i, Y_i) = aX_i + bY_i$. The random components were normally distributed with means of zero and variances of one and the ϵ_1 's and ϵ_2 's were 95 percent correlated (Horowitz, 1981, p 431-32). It should be reiterated that the probit model allows for correlation between the errors of different alternatives and therefore, IID is not as restrictive for the probit model (see section 2.2). The standard deviation of the differences between the probit and logit probability predictions equaled 0.022 with a maximum error of 0.074. Furthermore, Horowitz investigated the stability of the logit model predictions in conditions where the random utility components had unequal variances. The model performed favourably with a root mean square error of 0.027 and a maximum error of 0.13 (Horowitz, 1981, p 431). It would therefore appear that the concern with the IID assumption may not be a critical issue.

2.7.3 Measurement Error

Measurement error is a negligible source of error in the estimation of the MNL. As Hensher and Johnson (1981) suggest, "since the dependent variable is quantal, we have no measurement error from this variable" (p 224). Consequently, measurement error is restricted to the occasions where group means are used for explanatory variables (Horowitz, 1981, p 429) and the rare situation in which individuals are compelled to indicate their position in socio-economic categories and perform this task incorrectly (Hensher and Johnson, 1981, p 225). The only exception is recording errors by researchers.

Horowitz (1981) appears to have been the only author to address this problem in respect to discrete choice models. He investigated the use of "group-mean values of explanatory variables" though he contends that this is a special case of omitted variable bias where the omitted variable is the difference between the individual and group mean value (Horowitz, 1981, p 429). This type of error will lead to inconsistent estimates of choice probabilities for individuals and also at the group mean level. However, consistent estimates will be produced in two circumstances. First, where the group explanatory variables are not correlated with any nongrouped variables that are included in the specification. And second, "when the grouped variables have the same joint distribution function in each group, both in the estimation data set and the set used for forecasting" (Horowitz, 1981, p 429).

In his numerical example, Horowitz compared a correctly specified three alternative logit model with one where the data had been grouped into sets of five and the group mean value was used. The effect on the choice probabilities was considerable with a root mean square absolute error of 0.11 and a maximum error of 0.35 which is significant when compared to other errors investigated by Horowitz (1981, p 431). Again, scale parameter differences, not accounted for by Horowitz, could have produced this result.

2.7.4 Aggregation Bias

Aggregation is driven by the need for decision-makers to make predictions about the behaviour of groups of individuals in response to different marketing variables under their control, rather than individuals per se (Hensher and Johnson, 1981, p 194; Ben-Akiva and Lerman, 1989, p 131). Instead of modelling the behaviour of individuals, it is more judicious in terms of prediction to model their aggregate behaviour. However, the process of aggregation can lead to bias in the estimates and forecasts produced by the model. In this context, aggregation bias has a well defined meaning:

“The correct approach is to use the values of the independent variables that are directly relevant to each individual in the affected population, to predict the individual probabilities, and then to sum these probabilities. The predictions obtained by the individual disaggregative approach ... [the enumeration method]... will be more accurate than those obtained by the aggregate approach ...[naive aggregation method] ... because the average of a non-linear function is not equal to the function evaluated at the average of the explanatory variables. The difference between the predictions obtained by the enumeration method and the naive method is referred to as the aggregation bias” (Hensher and Johnson, 1981, p 195).

Aggregation bias occurs at a number of levels but Hensher and Johnson (1981) underscore the importance of two: the application of group and between group averages to bound the dimensions of the independent variables for each individual, and the use of the average individual level of a variable to obtain choice probabilities “both before and after a policy change” (pp 196-197).

The first type of error is “minimal” according to the research by McFadden and Reid (1975) if a random sample is used and the effect being measured is homogeneous within a group.

Furthermore, Koppelman (1975, 1976, 1976a) following his extensive research into the second

type of aggregation bias concluded that in practice the size of the bias was small compared to other sources of error. Nevertheless, Hensher and Johnson (1981) advise that these findings should be accepted with caution given the limited nature of the respective investigations.

Hensher and Johnson (1981) suggest that the most flexible method for overcoming aggregation problems associated with the “non-linearity of the choice function” is to use the sample emuneration method:

“This approach involves:

- (i) predicting the behaviour of each individual in a sample drawn from the population and taking the average of those predictions;
- (ii) predicting the after behaviour for each individual and taking the average;
and
- (iii) identifying the differences between those averaged predictions to obtain the aggregate policy effect” (Hensher and Johnson, 1981, p 199).

The main problem with this approach is the prohibitive costs involved. Other methods which exist involve approximation of the empirical distribution through the use of a mathematical distribution, a histogram, and Taylor series expansions of the statistical moments.

2.7.5 Systematic Sampling Error

Systematic sampling error or bias is an outcome of defects in the process of sample selection. This type of error is the principal focus of this study. Although specification and measurement error as well as aggregation bias are other sources of systematic sampling error (Hensher and Johnson, 1981, p 225), they are treated separately in this research. Hensher and Johnson (1981) summarise the conventional sources of systematic sampling error:

- (a) deliberate selection of an average sample;
- (b) sampling on the basis of an attribute that is correlated with one or more properties of the observational unit;
- (c) selection of a random sample in which the random selection process is not strictly adhered to;
- (d) the substitution of additional members of the population when difficulties are encountered in sampling the original sample observation; and
- (e) nonresponse (p 225).

All of these are widespread in survey research with the final three sources becoming more prevalent. Surprisingly, there has apparently been little research into the effect of systematic sampling error or bias on the estimates produced by the logit model.

2.7.6 Random Sampling Error

Random sampling error reflects the probabilistic nature of the sample selection procedure. It is an indication of the random differences between the sample and the population from which it was drawn. "Random sampling error (rse) is, all other things being constant, approximately proportional to the inverse of the square root of the sample size" (Hensher and Johnson, 1981, p 226). Random sampling error is a consequence of sample size and population variance (Hensher and Johnson, 1981, p 226) and can be reduced by either increasing the sample size or through exploiting the variance characteristics of groups of individuals in the population.

The bootstrap research of Gordon, Lin, Osberg, and Phipps (1994) looked specifically at the issue of random sampling error. In particular, they investigated two important issues associated with the application of dichotomous logit and probit choice models, namely instability in the estimation of parameters and standard errors, and variation between the quantitative results generated by these different models. They used an existing data set of 18,350 observations to construct a model of female labour force participation with 22 independent variables and the dependent variable being the dichotomy participant/not a participant.

Random subsets of observations converging toward the full sample size were drawn from the complete data set and a model fitted to these. The resultant paths of convergence for one of the independent variables demonstrated that probit parameter estimates were not stable until sample sizes of over 10,000 were attained while the standard errors approached a stable level after 4000 observations (Gordon et al, 1994, p19-20). This analysis was not undertaken for logit.

Further to this, Gordon et al (1994) examined stability by computing for logit, probit, and OLS models the number of times out of twenty independent runs that the estimated parameter value lay within the 95% confidence interval of the full sample-value of the parameter. They found for all three methods that "samples of 9,000 to 10,000 observations are required before the researcher

can be reasonably certain of obtaining coefficient estimates within a 95 percent confidence interval around the full-sample estimates” (p24).

In comparing the logit and probit models, Gordon et al (1994) compared the marginal effects of a given change in the independent variable of interest on the predicted probabilities of labour force participation (p26). The results showed the logit model predicted marginal effects 2.5 percent higher than the probit model for small sample sizes but as the sample increased over 10,000 observations, the difference increased to four percent (p 29).

This study is promising but requires further extension as only one of the independent variables has been analysed and it is unknown if the patterns can be generalised to the remaining variables. More importantly, the use of an actual data set confounds the generalisability of the results as effects such as interactions between the independent variables and misspecification have not been controlled for in the design. Monte Carlo simulation through the use of artificially generated data sets could resolve this issue. Nevertheless, the differences appear to be substantial given the sample sizes of most MNL studies.

2.7.7 Summary

There has been comparatively little research on the effect of the possible sources of error on the MNL estimates and probability forecasts. This is probably an acknowledgement that solutions to these problems are complicated and costly to implement in practice. The more general MNP model is rarely used in marketing mainly for these reasons. Surprisingly, of the six sources of error, systematic sampling error or bias has been the least researched. The performance of the MNL model in conditions of bias is important given that in marketing, it has been calibrated on samples drawn from the population of interest.

Furthermore, the stability of the MNL model when exposed to sample bias (as well as the other errors) is an important consideration because it provides some indication of the soundness or otherwise of the assumptions underlying it.

2.8 Monte Carlo Research

Monte Carlo or simulation studies investigate the relationship between actual or postured data and simulated data from an assumed model. The simulated data is generated by drawing random samples from a known population distribution. The Monte Carlo method serves two main purposes. First, the likeness of the real and simulated data serves as an indication of the “goodness-of-fit” or reliability of the model (Ripley, 1987, p 4). Second, the Monte Carlo method can be used to derive the sampling distribution of the test statistic in cases where the distribution in the population is known (Noreen, 1989, p 44).

“The key element in the Monte Carlo sampling method, as in conventional parametric methods, is the population that is specified in the null hypothesis” (Noreen, 1989, p 49). Populations that can be sampled include the normal, exponential, log, and uniform distributions. Monte Carlo is similar to bootstrapping in that random samples are drawn from a known distribution to yield a sampling distribution. The main difference is bootstrapping takes samples from an actual data set collected from the population and therefore imposes the distribution of the actual data set on the subsequent sampling distribution.

Monte Carlo methods are usually employed to address mathematical problems that have proved intractable and unmanageable to other forms of analysis. These methods have been used to analyse the bias and variance properties of estimates, the power of different tests, compute unbiased predictors in nonlinear models and a number of other inferential statistical applications (Gourieroux and Monfort, 1993, pp 5-6). In the field of discrete choice analysis, the Monte Carlo method has been employed on a number of occasions (Chen and Anderson, 1993; Horowitz and Louviere, 1993; Bunch and Batsell, 1989; Johnson, Meyer, and Ghose, 1989; Horowitz, 1981; Wittink and Cattin, 1981).

Chen and Anderson (1993) used Monte Carlo to investigate the small sample properties of the

logit maximum likelihood estimator. The utility of each alternative was generated by randomly drawing values of the parameters of the independent variables from a normal distribution, summing these values, and adding an exponential deviate (Chen and Anderson, 1993, p 243). Then, the individual response was determined for each choice set by assigning a value of 1 to the alternative with the highest utility for each individual and 0 to the other alternatives. Their reasons for sampling in this way are not clear but may be due to the fact that they were trying to control the value of the scale parameter for each model, and this process allowed this.

The MNL is frequently evaluated by using the estimates it makes to predict the choices in a holdout sample. Horowitz and Louviere (1993) used Monte Carlo to examine the finite-sample properties of a test they developed to compare predicted and observed choices (p 271). They generated estimation and test data sets "...by randomly sampling from a specified MNL model" for sample sizes of 100, 250, and 500 and repeating the process 1000 times to gain estimates of the size of the tests (Horowitz and Louviere, 1993, p 275). They concentrate on drawing random values of parameters rather than random individuals from a known distribution.

Bunch and Batsell (1989) used Monte Carlo to determine the relative precision of the estimates produced by the maximum likelihood, minimum chi square, minimum Pearson chi square, nonlinear least squares, and ordinary least squares estimators (p 57). Two factors were varied in a controlled fashion in the generation of the data: the scale parameter (40 percent explained variance and 70 percent explained variance) and the replication factor representing the number of responses in each cell (10, 40, and 80). "Repeated selections were simulated by first calculating the true choice probabilities for each subset and then drawing multinomial random vectors with the appropriate replication factor r " (Bunch and Batsell, 1989, p 60). The reasons for this distribution and method are not explained.

Johnson, Meyer, and Ghose (1989) examined the performance of the compensatory MNL model in non-compensatory environments. A number of different decision rules were simulated including elimination by aspects, lexicographic, conjunctive, phased EBA/compensatory, random,

and additive compensatory. Of particular interest to this study is the last decision rule. Individual responses in this situation were simulated by drawing samples from a normal distribution of parameter estimates. Furthermore, Johnson, Meyer, and Ghose (1989) introduced variance in the choices by adding random noise to the parameters on each trial (p 258). This was achieved by sampling from a normal distribution with a mean of zero and a variance of 0.2. "Our simulations therefore can be thought of representing the choices made by a homogenous segment of decision makers whose decision criteria contain a reasonable and realistic amount of random error" (Johnson, Meyer, and Ghose, 1989, p 259). This study was one of the few to explicitly discuss the nature of their method.

In the most comprehensive Monte Carlo study to date, Horowitz (1981) explored the impact of sampling, specification and data errors on the choice probability estimates produced by logit models. He generated a data set of 1000 observations for each type of problem investigated though the method he utilises to generate the data set is not clear. Furthermore, Horowitz (1981) does not seem to account for the scale parameter which could confound his results.

Wittink and Cattin (1981) use Monte Carlo to assess the predictive validity of models fitted to data generated under varying conditions. They used preference rank order data at the individual level in their investigation which is not entirely legitimate as the logit performs best on choice based data at the aggregate level. Again, their method for generating data is not lucid.

3.0 OBJECTIVES

The overall objective of this research was to fill the apparent gap in the existing modelling literature by examining the effect of systematic sampling error (or bias) on the parameter estimates made by the multinomial logit (MNL) choice model. At issue is the sensitivity of the logit model to this form of error and consequently, the generalisability of the estimates it makes to other similar populations. In marketing, it is customary for the MNL to be fitted to samples. It would be desirable, for practitioners and academics alike, if in the presence of sample bias, the MNL model produced estimates that were strongly representative of the effects being measured. Evidence of this nature would support the MNL's calibration on moderately biased samples, as the researcher would be confident that the effect being measured was being reflected in the parameter estimates more than the error. Furthermore, the assumptions underlying the MNL would to some extent be corroborated, thus providing tentative support of the robustness of the model in predicting choice.

In this study, an insensitive model is one whose estimates reflect the effect being measured rather than the error or bias present in the data. A sensitive model is one which in the presence of error produces estimates that reflect that error more than the effect. For example, a binary logit model ($n=100$) estimated on an error-free sample produces the following estimate $V(X_i, Y_i) = 10 X_i + 5Y_i$. A sensitive model estimated on a biased sample, such that the ten consumers most responsive to variable X are missing, would yield, for example, the estimate $V(X_i, Y_i) = 6 X_i + 5Y_i$ (i.e. the inability to sample the highest ten percent of the population on variable X produced a greater than ten percent change in the parameter estimate of X). Conversely, an insensitive model estimated on this same sample might produce the parameters $V(X_i, Y_i) = 9.9 X_i + 5Y_i$.

Furthermore, if the sample is biased on one independent variable (as in the previous example where bias occurred on X), then it would be desirable for the estimation of other effects (such as Y above) to occur independently of this. That is, the estimate of Y should not be influenced by

the bias on X . If it is, it indicates further instability in the model. It is not clear, a priori, whether this independence should hold for the MNL model, nor under what circumstances.

The specific objectives of this study were:

1. To verify that the Monte Carlo method used in this study recovered the original or 'true' parameter values by comparing these original estimates with the expected values of the unbiased sampling distribution:

H_{01} The expected values of the coefficients in the unbiased sampling distribution equal the original (or 'true') coefficients, and;

H_{11} The expected values of the coefficients in the unbiased sampling distribution are significantly different from the original coefficients at the 95 percent level of confidence.

2. To determine if the effect of biasing a sample by a fixed proportion on one variable created the same proportional change in that variable's coefficient estimates:

H_{02} The change in the normalised expected value of the biased variable from the unbiased to the biased sampling distribution equals the proportion of sample bias being simulated, and;

H_{12} The change in the normalised expected value of the biased variable from the unbiased to the biased sampling distribution is greater than the proportion of sample bias being simulated at the 95 percent level of confidence.

3. To determine if biasing a sample by a fixed proportion on one variable produced no change in the coefficient estimates of the other independent variables:

H_{03} The change in the normalised expected values of the variables upon which sample bias *did not* occur, from their value in the unbiased to the biased sampling distribution equals zero, and;

H_{13} The change in the normalised expected values of the variables upon which sample bias *did not* occur from the unbiased to the biased sampling distribution is significant at the 95 percent level of confidence.

4. To ascertain if the ratio between the unbiased independent variables is not changed by the simulated sample bias. Stated differently, this tests if some variables are more sensitive to the simulated sample bias than others:

H_{04} The change in the normalised expected values of the variables upon which sample bias *did not* occur is not significantly different from other equivalent variables, and;

H_{14} Any change in the normalised expected values of the variables upon which sample bias *did not* occur is significantly different from other equivalent variables at the 95 percent level of confidence.

And, as an aside to the main issue of sensitivity of the MNL coefficients to bias:

5. To establish the accuracy of the sampling errors predicted by the MNL estimation software by comparing them with the standard deviation of the unbiased sampling distribution.

4.0 METHOD

This study used the Monte Carlo method to examine the behaviour of the MNL model when calibrated on biased samples. Specifically, the impact on the parameter estimates of removing varying proportions (five, ten and fifteen percent) of simulated individuals from both tails of the sample distribution was investigated. The extent to which marketing researchers fail to sample individuals at the extremities is unknown. Nevertheless, examining the variability of the MNL estimates in these conditions, arguably, provides a more rigorous test of the model's stability than removing individuals nearer the centre of the distribution.

The Monte Carlo method produces a sampling distribution for a particular statistic by drawing a sufficient number of random samples from a defined population with a specified distribution. In this case, the population is defined in terms of the parameter estimates gained from a 'true' MNL model of vacation destination choice behaviour. This true model was derived by calibrating the MNL on a sample of 164 individual's aggregated responses to the twelve dichotomous choice sets (see section 4.1 for more details on the design). The model calculated the effect of nine attributes of vacation destination and the intercept on choice.

Consequently, these true estimates were used in combination with the MNL to calculate the probability that the first alternative in each choice set in the original design would be selected. These probabilities were retained as the 'true' likelihood of a simulated individual selecting the first alternative. A computer programme was written in Fortran 77 to generate samples of simulated individuals.

For each simulated individual, twelve random (uniform) numbers were produced representing that individual's response to the twelve choice sets. If the random number was less than or equal to the true likelihood for the first alternative, then the individual was deemed to have selected the first alternative. Conversely, if the random number was greater than the true likelihood, then the

second alternative was judged to have been chosen. To produce one simulated individual, this process was performed for each of the twelve choice sets and twelve respective random numbers. Each simulated individual is produced so that the interaction between the attributes equals zero. This is because this study is primarily interested in the impact of bias on the nine main effects rather than any interactions between them. Consequently, constraining the interactions to zero controls for any possible confounding effect they may have on the results.

Initially, an unbiased sampling distribution of the MNL coefficients was created. The expected values of this sampling distribution served two purposes. First, they were compared with the true estimates used to produce the simulated individuals. This acts as a guide to the Monte Carlo's ability to recover the true population values. Second, these expected values were the benchmark to which the expected values in biased sampling distributions could be compared. Furthermore, the characteristics of the unbiased sampling distribution as expressed by the kurtosis, skewness, and range were used to test the normality of the estimates. This is important as the estimation software used assumes that the coefficients are normally distributed when computing the standard errors.

The unbiased sampling distribution was produced by repeatedly fitting the MNL to randomly generated unbiased samples of 164 simulated individuals. The individuals were created as described above and their responses to the twelve choice sets were aggregated. The MNL was estimated by regressing these aggregate responses with the design matrix using the NTELOGIT estimation programme. In total, the MNL was fitted to 10,000 samples of 164 individuals.

Biased sampling distributions of the MNL coefficients were created for each type of sample bias under investigation. Sample bias was simulated by eliminating those individuals whose responses were most (or least) highly correlated with one of the nine explanatory variables (the intercept is not included). This resulted in 54 different sampling distributions being produced (bias on nine variables at the tails of the distribution at three levels: $9 \times 2 \times 3 = 54$). The biased sampling distributions were created by repeatedly fitting the MNL to randomly generated biased samples of

164 simulated individuals. A biased sample was created by generating 173, 182, and 193 simulated individuals for five, ten, and fifteen percent bias respectively, then removing those who were most (or least) highly correlated with one of the nine explanatory until 164 individuals remained.

In this study, the MNL is always fitted to samples containing 164 simulated individuals. This permits a direct comparison of the various biased sampling distributions with the benchmark unbiased sampling distribution. Moreover, it enables a comparison of the parameter estimates and standard errors in the original data set of 164 individuals with the expected values and standard deviations in the unbiased sampling distribution (objectives one and five).

Two other methods could have been used to address the objectives of this study, namely calculus and bootstrapping. A review of the research revealed that mathematically solving problems relevant to the MNL had only been undertaken for limited independent variables (usually two) and for dichotomous choice. For a problem of this size with nine independent variables and a variety of different methods for biasing the sample, the Monte Carlo approach was viewed as offering a more flexible procedure. Bootstrapping involves drawing random sub-samples from an actual set of data to derive the theoretical sampling distribution. It was rejected as a method for deriving the unbiased and biased sampling distributions as the results would have been confounded with other factors such as interactions between the independent variables and as such, the generalisability of the results of this type of study would have been questionable.

The following sections outline in more detail than above, the original design and true model, the method used to generate individuals and samples, the procedure used to simulate sample bias, the procedure used to produce the unbiased and biased sampling distributions, the Fortran 77 programme, and the scale parameter normalisation.

4.1 Experimental Design and Calibrated MNL Model

To produce results that were relevant to the MNL model's application in a marketing environment, the results of a study on destination choice were used to generate the required sampling distributions. This study had examined nine variables and their impact on the choice of vacation destination. These variables included:

- total cost of the vacation;
- the distance to travel to the vacation destination;
- the number of activities available at the destination that the respondent likes to do;
- weather and climate;
- the need for concern about food and/or water;
- the ease of undertaking the vacation without learning a new language;
- the beauty of the scenery;
- safety with regard to crime and terrorism, and;
- whether the destination had new or different things to see and experience.

A 2^9 fractional factorial, main effects design was used by the original researchers to estimate the coefficients of these nine variables (see table 1 for the design). All main effects were balanced across the design. The fact that main effects alone and not interactions were examined in this research is not viewed as being to the detriment of the results and their generalisability. Although one of the reasons given for measuring the values of the interactions is to reduce the level of bias in the parameter estimates (Louviere, 1988, p 40), the method used to generate individuals in this study means that all interactions should be equal to zero. Nevertheless, using a design that estimates the interactions as well as the main effects may reduce the overall sensitivity of the MNL model to bias. The extent to which this could be the case is unknown.

Table 1. Design of the 12 Choice Sets in the 2⁹ Factorial Main Effects Vacation Destination Choice Experiment.

Choice Set/ Alternative	β_1 How much is the total cost of a vacation	β_2 How long is the trip to get there	β_3 How many things I like to do can I do	β_4 How good /bad is the weather/ climate	β_5 Need for concern about food and/or water	β_6 Ease of getting by without a new language	β_7 How much beautiful scenery is there	β_8 Is it safe from crime /terrorism	β_9 New/ different things to see /experience
1 A	LOW	LONG	MANY	BAD	LITTLE	HARD	MUCH	UNSAFE	MANY
1 B	HIGH	SHORT	FEW	GOOD	MUCH	EASY	LITTLE	SAFE	FEW
2 A	LOW	SHORT	MANY	BAD	LITTLE	EASY	LITTLE	SAFE	MANY
2 B	HIGH	LONG	FEW	GOOD	MUCH	HARD	MUCH	UNSAFE	FEW
3 A	HIGH	LONG	FEW	BAD	LITTLE	HARD	LITTLE	SAFE	FEW
3 B	LOW	SHORT	MANY	GOOD	MUCH	EASY	MUCH	UNSAFE	MANY
4 A	HIGH	SHORT	MANY	GOOD	LITTLE	HARD	MUCH	SAFE	FEW
4 B	LOW	LONG	FEW	BAD	MUCH	EASY	LITTLE	UNSAFE	MANY
5 A	LOW	SHORT	FEW	BAD	MUCH	HARD	MUCH	UNSAFE	FEW
5 B	HIGH	LONG	MANY	GOOD	LITTLE	EASY	LITTLE	SAFE	MANY
6 A	HIGH	LONG	FEW	GOOD	LITTLE	EASY	MUCH	UNSAFE	MANY
6 B	LOW	SHORT	MANY	BAD	MUCH	HARD	LITTLE	SAFE	FEW
7 A	HIGH	SHORT	MANY	GOOD	MUCH	HARD	LITTLE	UNSAFE	MANY
7 B	LOW	LONG	FEW	BAD	LITTLE	EASY	MUCH	SAFE	FEW
8 A	HIGH	SHORT	FEW	BAD	MUCH	EASY	MUCH	SAFE	MANY
8 B	LOW	LONG	MANY	GOOD	LITTLE	HARD	LITTLE	UNSAFE	FEW
9 A	LOW	SHORT	FEW	GOOD	LITTLE	EASY	LITTLE	UNSAFE	FEW
9 B	HIGH	LONG	MANY	BAD	MUCH	HARD	MUCH	SAFE	MANY
10 A	LOW	LONG	FEW	GOOD	MUCH	HARD	LITTLE	SAFE	MANY
10 B	HIGH	SHORT	MANY	BAD	LITTLE	EASY	MUCH	UNSAFE	FEW
11 A	LOW	LONG	MANY	GOOD	MUCH	EASY	MUCH	SAFE	FEW
11 B	HIGH	SHORT	FEW	BAD	LITTLE	HARD	LITTLE	UNSAFE	MANY
12 A	HIGH	LONG	MANY	BAD	MUCH	EASY	LITTLE	UNSAFE	FEW
12 B	LOW	SHORT	FEW	GOOD	LITTLE	HARD	MUCH	SAFE	MANY

In the original study, a total of 164 individuals responded to all twelve of the dichotomous choice situations. The data set consisted of their aggregate responses. These were used to estimate a logit model by regressing their aggregated responses against the design matrix using the NTELOGIT estimation programme. This yielded the coefficient estimates in table 2. Variable seven (scenery) has the greatest impact on vacation destination choice than the other variables; cost (variable one) has the least impact. The presence of two variables has a negative impact on vacation destination choice, variable one (cost) and variable nine (newtodo). It is not surprising that cost decreases the chances of choosing a particular destination, but it is unexpected that “a location with more new activities” has less chance of being selected. The asymptotic t-values for the nine independent variables and the intercept indicate that the hypothesis that the coefficients are equal to zero can be rejected at the 95 percent level of confidence. With an adjusted rho-squared of .91, this model has an excellent fit to the data and is a more than adequate representation of choice in this market.

Table 2. MNL Model of Vacation Choice Behaviour Estimated on the Original Set of 164 Individuals

VARIABLE	COEFFICIENT	STD ERROR	T
β_7 scenery	0.965684	0.071767	13.4558
β_2 long	0.880865	0.070605	12.4759
β_3 todo	0.874663	0.067876	12.8861
β_4 climate	0.599239	0.070007	8.5597
β_8 safety	0.466055	0.071382	6.5290
β_6 language	0.429348	0.071921	5.9697
β_9 newtodo	-0.312841	0.060666	-5.1568
β_5 food	0.269242	0.069957	3.8487
β_0 intercpt	0.231840	0.067090	3.4557
β_1 cost	-0.130643	0.064260	-2.0330
L(ZERO):	-472.49	L(BETA):	-30.49
-2(L(0)-L(B)):	884.00	D.F.:	10
RHOSQ:	.93546	ADJUSTED RHOSQ:	.91430

The parameter estimates gained from this aggregate MNL model were treated as the 'true' values for vacation destination choice behaviour in the Monte Carlo study and were retained for use in the generation of individual level responses. In effect, we are assuming that the MNL is the true model of vacation choice behaviour. The main consequence of this is that the sampling distributions produced in this thesis represent a group of consumers with homogenous MNL preferences. This is restrictive in that most consumer markets consist of individuals or groups with heterogenous preferences. Therefore, the potential for sampling bias is greater if the heterogeneity is not accounted for. As a result, if the MNL is found in this study to be relatively insensitive to bias, this does not mean that this is also the case for heterogenous markets. However, if the MNL is found to be sensitive to sampling bias in this instance, then it is likely to be more sensitive when calibrated in heterogenous markets.

4.2 Generating Individuals and Samples

After fitting the MNL model to the original set of 164 individuals' aggregated responses and assuming that the subsequent estimates were true, a method for producing simulated individuals and samples was required. This was accomplished by using the original estimates (table 2) and the MNL to calculate the probability that the first of the two alternatives in each of the twelve choice sets in the original design would be selected. These probabilities or 'true' likelihoods were then used as 'cutoff' values in the generation of random individuals. The cutoff values for the twelve choice sets were:

Choice Set One	=	.66372888;	Choice Set Two	=	.92773445;
Choice Set Three	=	.17265963;	Choice Set Four	=	.95538925;
Choice Set Five	=	.04968568;	Choice Set Six	=	.64114886;
Choice Set Seven	=	.50192799;	Choice Set Eight	=	.45641192;
Choice Set Nine	=	.41982691;	Choice Set Ten	=	.63221849;
Choice Set Eleven	=	.94095225;	Choice Set Twelve	=	.08343799.

For each new individual, 12 random numbers ($0 \leq X \leq 1$) were generated and represented that individual's response to the twelve choice sets. If the first random number was less than or equal to the cutoff for choice set one, then the simulated individual was treated as having selected the first alternative. Conversely, if the first random number was greater than the cutoff for choice set one, then the simulated individual was treated as having selected the second alternative. For each individual, this was continued for all remaining eleven respective choice sets using the random numbers and cutoffs. The result was a simulated individual and their responses to the twelve choice sets.

This process is analogous to individuals whose responses were drawn from a Bernoulli distribution: "This serves as a model for any situation in which we can think of a trial being made with probability p of 'success' and probability $(1 - p)$ of 'failure' " (Cooke, Craven, and

Clarke, 1990, p 109). This would seem appropriate for the process being modeled in this case; as we assumed the MNL to be the true model of vacation behaviour, the individual choice of an alternative in a choice set depended upon a predetermined probability of success and failure. The use of this Bernoulli distribution should not impact upon the distribution of the MNL coefficients as the maximum likelihood estimator produces asymptotically normal estimates regardless of the distribution being sampled (Ramanathan, 1995, p 76).

The number of individuals generated depended on the type of bias being simulated. For the unbiased sampling distribution, samples of 164 were required. For five percent bias, samples of 173 were required so that nine individuals could be removed from the sample in order to simulate sample bias. For ten percent bias, samples of 182 were needed so that 18 people could be deleted. For fifteen percent, samples of 193 were generated so that 29 individuals could be removed.

4.3 Biasing the Sample

The individuals were removed from the sample on the basis of their 'score' on one of the independent variables. An individual with a 'high score' was someone whose choice was highly correlated with the particular independent variable. An individual with a 'low score' was the opposite and is analogous to someone whose choice is not influenced by that particular independent variable. As the experimental design was perfectly balanced, each effect appeared in the design 24 times, 12 times for each level (these levels were comparable to the presence or absence of the effect). Consequently, an individual's 'score' was determined by calculating the number of times a respondent had selected an alternative that contained the independent variable upon which the sample was being biased when that variable was 'present'.

Individuals were then removed from the sample on the basis of their score. If bias at the top of the distribution was being simulated, then individuals were first removed if they had a score of twelve until the required number had been deleted. If the number of people with a score of twelve was less than the number that had to be removed, then individuals were then removed if they had a score of eleven. This process was continued until the desired number of people had been deleted.

4.4 Generating the Unbiased Sampling Distribution

The unbiased sampling distribution was generated by fitting the MNL to 10,000 random samples of 164 simulated individuals. This involved:

1. Calculating for the first alternative in the twelve choice sets, the probability that that alternative would be chosen:

$$P(A \in S_n) = \frac{e^{\beta \cdot x_A}}{e^{\beta \cdot x_A} + e^{\beta \cdot x_B}}$$

where: x_A = the attribute descriptions of the first alternative
 x_B = the attribute descriptions of the second alternative
 β = the vector of true MNL regression coefficients.

2. Generating a 164 X 12 matrix ($RAND_{JK}$) of random numbers (R_n) to represent each individual's response to the twelve choice sets ($0 < R_n < 1$).
3. Converting the matrix of random numbers above into a discrete format (0,1) where 1 is equivalent to a positive response, or selection of the first alternative (and hence, a negative response to the second alternative), and 0 represents non-selection of the first alternative (and therefore, a positive response to the second):

If $RAND_{JK} \leq P(A \in S_n)$ then $RESP_{JK} = 1$

If $RAND_{JK} > P(A \in S_n)$ then $RESP_{JK} = 0$.

Where : $J = 1$ to 164

$K = 1$ to 12

4. The individuals' responses to each choice set were aggregated:

$$\text{AGGRES}_{JK} = \sum_{K=1}^{164} \text{RESP}_{JK}$$

5. This aggregate matrix was combined with the design matrix and the NTELOGIT estimation yielded the unbiased estimates of β . These ten estimates were then written to file.
6. This process was repeated 10,000 times to yield an unbiased sampling distribution of the complete MNL model (nine independent variables plus the intercept).

4.5 Generating the Biased Sampling Distribution

The biased sampling distributions ($n = 10,000$) were generated using a similar procedure:

1. Calculating for the first alternative in the twelve choice sets, the probability that that alternative would be chosen:

$$P(A \in S_n) = \frac{e^{\beta \cdot x_A}}{e^{\beta \cdot x_A} + e^{\beta \cdot x_B}}$$

- where: x_A = the attribute descriptions of the first alternative
 x_B = the attribute descriptions of the second alternative
 β = the vector of true MNL regression coefficients.

2. Generating a $T \times 12$ matrix ($RAND_{JK}$) of random numbers (R_n) to represent each simulated individual's response to the twelve choice sets ($0 < R_n < 1$). The size of T depended upon the level of bias being investigated. In this case, T equaled:

- 173 when the level of bias was five percent;
- 182 when the level of bias was ten percent, and;
- 193 when the level of bias was fifteen percent.

3. Converting the matrix of random numbers above into a discrete format (0,1) where 1 is equivalent to a positive response, or selection of the first alternative (and hence, a negative response to the second), and 0 represents non-selection of the second alternative (and therefore, a positive response to the first):

$$\text{If } RAND_{JK} \leq P(A \in S_n) \text{ then } RESP_{JK} = 1$$

If $RAND_{JK} > P(A \in S_n)$ then $RESP_{JK} = 0$.

Where : $J = 1$ to T

$K = 1$ to 12

4. Removing $T - 164$ simulated individuals from the matrix $RESP_{JK}$. For each individual, create a new variable (SCORE) which equals the number of times they selected an alternative that contained the independent variable (attribute) upon which the sample was to be biased. This was achieved by comparing their response to the first alternative with the design of that alternative. If they selected the first alternative ($RESP_{JK}$ equals 1) and the alternative contained the bias variable at the required level ($DESIGN_{LM}$ equals 1), then one would be added to their score. Alternatively, if their response to alternative A was negative ($RESP_{JK}$ equals 0) and alternative A didn't contain the bias variable at the required level ($DESIGN_{JK}$ equals 0), then as the design was balanced and the choice a dichotomy, this was treated as equivalent to that individual selecting the bias variable in the second alternative, so one was added to their score.

If $RESP_{JK} = DESIGN_{LM}$ then $SCORE_{JK} = SCORE_{JK} + 1$

Where : $J = 1$ to T

$K = 1$ to 12

$L = 1$ to 12

$M =$ bias variable (from 1 to 9)

$DESIGN$ is the design of the nine attributes of alternative A.

The individuals to be removed to simulate bias are then 'marked' so they are not aggregated in the $RESP$ matrix in the next step. This was achieved by changing each J in the matrix $RESP_{JK}$ to -1 for those individuals to be removed. The individuals to be deleted were determined by their score from above and the type of bias being simulated. If individuals were being removed from the top (bottom) of the distribution, then those T -

164 individuals with the highest (lowest) score were marked.

5. After biasing the sample, the remaining individuals' responses to each choice set were aggregated:

$$\text{AGGRES}_{JK} = \sum_{K=1}^{164} \text{RESP}_{JK}$$

6. This aggregate matrix of responses was added to the design matrix and the NTELOGIT estimation program regressed it against the design matrix to yield the unbiased estimates of β . These ten estimates were then written to a file.
7. This process was repeated 10,000 times to yield biased sampling distributions of the complete model (nine independent variables plus the intercept).

4.6 FORTRAN 77 Computer Programme

A computer programme was written in FORTRAN 77 to produce the MNL coefficient estimates required for the biased and unbiased sampling distributions. The Monte Carlo programme written was basically a shell around the existing FORTRAN 77 logit estimation programme called NTELOGIT. The main change to this was in the command input subroutine (SUBROUTINE CMDINP) which reads the information from a pre-written command file each time before it estimates the model. This command file sets the format and characteristics of the data and design matrices of the MNL model. As these variables were identical for the models estimated in this study, rather than having the estimation programme read the command file before each calibration, the characteristics were fixed within the estimation programme to decrease time requirements. The output was also adapted so that the only information being returned was the ten coefficient estimates.

The Monte Carlo programme consisted of two main subroutines. The first, SUBROUTINE GENERATE produced the 164×12 matrix of aggregated biased or unbiased responses to the first alternative in the twelve choices. The second SUBROUTINE ESTIMATE consisted of the modified NTELOGIT estimation programme. The other main feature of the programme is the FUNCTION RANDOM which is the Applied Statistics Algorithm AS183 (1982) that returns a psuedo-random number between zero and one.

The programme is in Appendix A and contains extensive comments explaining its structure.

4.7 Scale Parameter Normalisation

As mentioned in the literature review, the scale parameter has a direct impact on the size of the coefficients estimated by the MNL model. The scale parameter is inversely related to the error variance (Ben-Akiva and Lerman, 1989, p 104). As the variance approaches infinity, the scale parameter tends towards zero. Conversely, the scale parameter tends toward infinity as the variance approaches zero. The effect of this on the β 's is not a minor consideration. Large scale values generate large β 's and vice versa. Consequently, the scale of the β 's determines the amount of variance explained by the systematic component of the utility function relative to the random component.

The value of the scale parameter of the β 's is especially important when MNL models estimated on different samples are being compared (Swait and Louviere, 1993). This is because any observed disparity between the models' estimates could be as a result of scale parameter variation rather than actual sample choice differences. Therefore, it is reasonably clear that scale parameter differences have to be removed before any variation in the coefficients from separate samples can be examined.

In this study, the expected values of the ten coefficients in the unbiased and biased sampling distributions are normalised as follows:

$$\text{Normalised } E(\beta_{\text{biased}}) = E(\beta_{\text{biased}}) \times \frac{\sum_{u=1}^8 |E(\beta_u)|}{\sum_{b=1}^8 |E(\beta_b)| + \sum_{u=1}^8 |E(\beta_u)|}$$

where: $|E(\beta_b)|$ = the absolute value of the expected values of the coefficients in the biased sampling distribution;

$|E(\beta_u)|$ = the absolute value of the expected values of the coefficients in the unbiased sampling distribution,

and;

u and b = the coefficients in the unbiased and biased sampling distributions from 1 to 8, excluding the biased variable and the intercept from the ten estimated in the MNL model.

$$\text{Normalised } E(\beta_{\text{unbiased}}) = E(\beta_{\text{unbiased}}) \times \frac{\sum_{b=1}^8 |E(\beta_b)|}{\sum_{b=1}^8 |E(\beta_b)| + \sum_{u=1}^8 |E(\beta_u)|}$$

where: $|E(\beta_u)|$ = the absolute value of the expected value of the coefficients in the unbiased sampling distribution;

$|E(\beta_b)|$ = the absolute value of the expected value of the coefficients in the biased sampling distribution, and;

u and b = the coefficients in the unbiased and biased sampling distributions from 1 to 8, excluding the biased variable and the intercept from the ten estimated in the MNL model.

This normalisation procedure assumes that the expected values of the coefficients in the unbiased and biased sampling distributions are linearly related. Therefore, the scale parameter differences between the two sampling distributions can be controlled for by a multiplicative scalar, representing the slope of the best fit straight line to the combined coefficients. Consequently, to adjust for the scale differences between the unbiased and biased sampling distributions, the expected values of all the independent variables are adjusted for by the slope of a straight line fitted to the expected values of the coefficients of the unbiased variables.

The biased variable and the intercept were not included in the normalisation formula. The biased variable is excluded as the effect of the sample bias on its coefficient would be confounded with

that of the scale parameter. As a result, it would be unwise to introduce other potential errors in the normalisation procedure. The intercept is excluded on the basis of the research by Lee (1982) and Swait and Louviere (1993). Lee (1982) found the intercept was biased in conditions where a relevant variable had been omitted. Swait and Louviere (1993) also found that the relationship between the constants of two different data sets was not consistent with the other coefficients. The reason for this is connected with the role of the intercept in maintaining the aggregate shares of the alternatives (Swait and Louviere, 1993, p 311) rather than measuring any distinct effect in the choice situation. It should be noted, however, that for the normalisation of the original coefficients and the expected values of the unbiased sampling distribution (section 5.1), all unbiased variables were included.

This normalisation procedure is not ideal for a number of reasons. First, the use of the expected values of the sampling distributions may introduce error into the process. A better method would involve normalising each of the 10,000 estimates of the unbiased variables in the unbiased sampling distribution with regard to their value in the biased sampling distribution. The average of these 10,000 normalised coefficients could then be used in the analysis. This would reduce the error in the normalisation as each model's distinct scale parameter would be included in the calculation. Despite this, it is speculated that any error introduced here would not be substantial

Second, a more suitable method that returns consistent estimates of the relative scale parameter between two data sets was proposed by Swait and Louviere (1993). The assumption is made that the MNL model is the true model of choice behaviour and that the specification of the model is the same for both samples. The data and design matrices of the two samples are then concatenated and a joint MNL model estimated. Following from this, the scale parameter of the first sample is fixed to one while for the second sample, the scale is changed until the log likelihood of the joint model is maximised. When this has been achieved, a consistent estimate of the relative scale has been attained.

Neither of these methods for adjusting for the scale parameter are employed in this study due to

the complexity of calculation for the large sampling distributions used here. However, any replication should attempt to test the effectiveness of the normalisation function used here relative to these better methods.

5.0 RESULTS AND DISCUSSION

5.1 Recovery of the original parameter values

Initially, a comparison was made between the 'original' or true coefficients from the MNL model estimated on the original 164 individuals and the expected values of the same coefficients in the unbiased sampling distribution. These original coefficients were used to generate the unbiased Monte Carlo sampling distribution and subsequent biased sampling distributions. Therefore, if the Monte Carlo method used in this study has any empirical validity, it should be able to produce a sampling distribution of parameter estimates with a mean (or expected value) closely approximating the original or true coefficients used to construct it. If not, then the sample bias effects being simulated later could be confounded with the inherent variability of the Monte Carlo method.

As previously mentioned, the 'original' coefficients were obtained by estimating a MNL model on the aggregated set of 164 actual individuals' responses to the 12 choice sets. An unbiased sampling distribution of the coefficients ($n=10,000$) was then created by randomly generating a response to each of the 12 choice sets in the design (based on the original coefficients) for the 164 simulated individuals, aggregating these individual's responses, and fitting the MNL to the ensuing data and design matrices. The expected values of the unbiased sampling distribution produced by the Monte Carlo method and the original sample parameters are shown in table 3.

As the expected values of the unbiased sampling distribution and the original coefficients varied consistently and linearly, it was hypothesised that the deviation was a result of scale parameter differences (see section 4.7 for a discussion of scale parameter normalisation). The expected values and original coefficients were then normalised to account for the difference in the respective scale parameters and these are also shown in table 3. Ideally, the normalisation would

have been performed separately for each of the 10,000 coefficients in the unbiased sampling distribution relative to the original coefficients. This would have provided a more accurate estimate of the variation between the scale parameters as each of the 10,000 MNL model's coefficients have a different scale. However, and consistent with all scale parameter normalisation in this study, the expected values of the coefficients in the unbiased sampling distribution are used. The error introduced by this is conjectured to be smaller than the effects being measured.

Table 3. How well does the Monte Carlo method recover the original coefficients?

Independent Variable	Original Coefficient	Monte Carlo Expected Value	Normalised Coefficient	Normalised Expected Value	Difference in Normalised Estimates
0 Intercept	0.232	0.238	0.117	0.118	+0.001
1 Cost	-0.313	-0.317	-0.158	-0.157	+0.001
2 Long	-0.131	-0.137	-0.066	-0.068	-0.002
3 Todo	0.881	0.894	0.445	0.443	-0.002
4 Climate	0.875	0.893	0.442	0.442	+0.000
5 Food	0.599	0.610	0.302	0.302	0.000
6 Language	0.269	0.269	0.136	0.133	-0.003
7 Scenery	0.429	0.442	0.217	0.219	+0.002
8 Safety	0.966	0.985	0.487	0.488	+0.001
9 Newtodo	0.461	0.466	0.233	0.231	-0.002

No differences significant at the 95 percent level of confidence

In all cases, the discrepancy between the original coefficients and the normalised Monte Carlo expected values is minimal. Furthermore, for the normalised coefficients these differences are not significant at the 95 percent level of confidence so the null hypothesis H_{01} is confirmed. It could therefore be safely concluded that the mean coefficient point estimates produced by the Monte Carlo method have performed as expected (subject to a linear transformation) and any future

variance in their value can be attributed to the controlled experimental differences.

5.2 Impact of Sampling Bias upon the Parameter Estimates of the Biased Variable

The degree to which sampling bias affects the coefficient estimates of the biased variables was determined by comparing their normalised expected values in the unbiased and biased sampling distributions. Biased sampling distributions were created under conditions where each sample of individuals on which the MNL model was to be calibrated was biased on each of the nine independent variables at each of the levels of interest. As a result, a total of 54 sets of biased sampling distributions ($n = 10,000$) were available for comparison, each one representing a specific variable upon which the sample was biased (nine), a different condition (individuals removed from the top or bottom) of bias, and a distinct level (five, ten, or fifteen percent removed). Again, to maintain consistency in this study, all MNL models were fitted to samples of size 164.

Table 4 illustrates the percentage change in the normalised expected values of the parameter estimates caused by the various forms of bias simulated. In other words, to what extent did the coefficients deviate between their unbiased and biased values in the respective sampling distributions. This was calculated as follows:

$$\text{Percentage change} = \frac{N E(\beta_u) - N E(\beta_b)}{N E(\beta_u)}$$

where : $N E(\beta_u)$ is the normalised expected value of the biased variable in the unbiased sampling distribution, and;
 $N E(\beta_b)$ is the normalised expected value of the biased variable in the biased sampling distribution.

Table 4. How much does bias affect the value of the biased variable?

Coefficient of the variable on which the sample was biased	Top Five Percent	Bottom Five Percent	Type of Bias		Top Fifteen Percent	Bottom Fifteen Percent
			Top Ten Percent	Bottom Ten Percent		
1 Cost	-24.59 %	+25.35 %	-42.31 %	+45.00 %	-62.39 %	+62.03 %
2 Long	-62.12	+62.57	-109.38	+104.64	-150.58	+152.32
3 Todo	-8.95	+7.72	-16.16	+13.22	-22.22	+19.78
4 Climate	-9.40	+8.50	-16.50	+14.74	-22.45	+21.95
5 Food	-14.28	+13.38	-23.96	+23.53	-34.62	+33.15
6 Language	-33.90	+33.09	-56.68	+54.97	-79.26	+76.92
7 Scenery	-19.74	+18.54	-33.90	+31.66	-47.28	+46.31
8 Safety	-8.41	+8.00	-14.81	+13.54	-20.25	+19.64
9 Newtodo	-19.69	+17.02	-28.87	+30.04	-45.91	+41.86

All differences are significant at the 99 percent level of confidence

It would appear to be obvious that the estimates produced by the MNL are sensitive to bias. For the overwhelming majority of cases, the percentage change in the value of the parameters far exceeds the proportion of individuals removed from the sample. The null hypothesis H_{02} is not supported by these results. That is, when ten percent of the sample is removed, the expected change (ten percent) in the value of the coefficient is surpassed.

Also significant is the fact that all the coefficients move in the expected direction. When the individuals who are particularly influenced by one variable are removed, the parameter estimates would be expected to decline and vice versa. This is important as it proves that the MNL estimates behave predictably even if their movement is somewhat greater than expected.

Another interesting finding is the greater than 100 percent change in the coefficient of variable two when exposed to ten and fifteen percent bias. It not only increases two-fold in value, but when sample bias of ten and fifteen percent occurs at the bottom, the coefficient changes signs

from negative to positive. Additionally, there is a noticeable difference between the effect of removing individuals from the top from that of the bottom of the distribution. Despite a few exceptions, most notably for the coefficients with negative effects (variable one and two), deleting individuals from the top of the distribution has a greater effect than removing individuals from the bottom. This difference could not be explained by the standard errors, the kurtosis, or the skewness of the sampling distributions of the biased coefficients.

Of most importance is the finding that the effect of the sampling bias depends upon the size of the respective parameters. Table 5 is a duplicate of the previous table except that the variables are ordered by the size of their coefficient (from largest to smallest). A plot of size of the original coefficients with the percentage effect of bias is shown in figure 1.

Table 5. Does the affect of the bias depend upon the value of the coefficient of the biased variable?

Coefficient of the variable on which the sample was biased	Top Five Percent	Bottom Five Percent	<u>Type of Bias</u>			
			Top Ten Percent	Bottom Ten Percent	Top Fifteen Percent	Bottom Fifteen Percent
8 Safety	-8.41%	+8.00%	-14.81%	+13.54%	-20.25%	+19.64%
3 Todo	-8.95	+7.72	-16.16	+13.22	-22.22	+19.78
4 Climate	-9.40	+8.50	-16.50	+14.74	-22.45	+21.95
5 Food	-14.28	+13.38	-23.96	+23.53	-34.62	+33.15
9 Newtodo	-19.69	+17.02	-28.87	+30.04	-45.91	+41.86
7 Scenery	-19.74	+18.54	-33.90	+31.66	-47.28	+46.31
1 Cost	-24.59	+25.35	-42.31	+45.00	-62.39	+62.03
6 Language	-33.90	+33.09	-56.68	+54.97	-79.26	+76.92
2 Long	-62.12	+62.57	-109.38	+104.64	-150.58	+152.32

All differences are significant at the 99 percent level of confidence

Figure 1. Plot of Coefficient with Percentage Effect of Bias

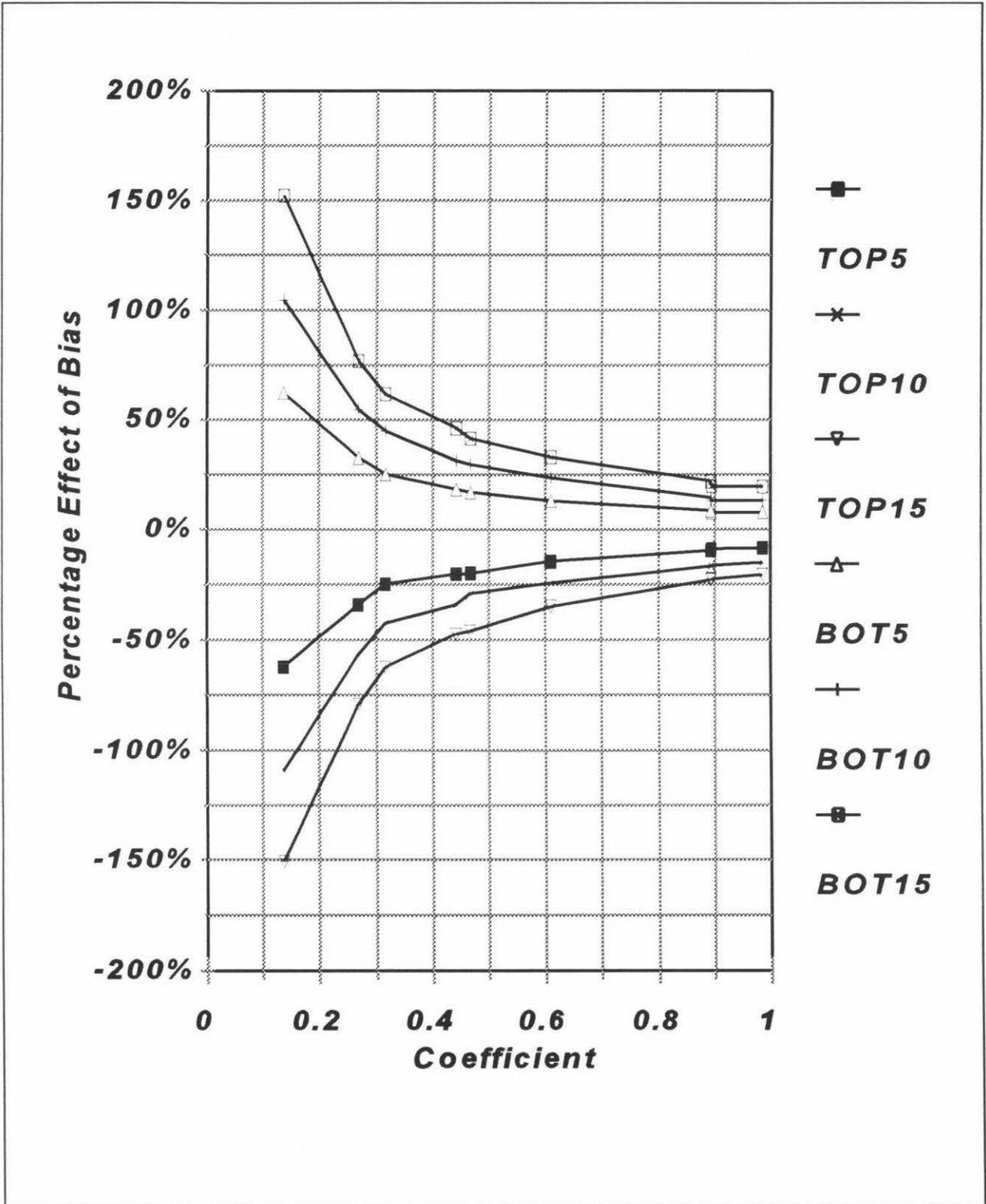


Table 5 and figure 1 above demonstrate that the proportional amount of change in the parameter estimates would appear to be explained by the relative value of the coefficient. A subsidiary null and alternative hypothesis is therefore developed as a result of this unexpected finding:

H_{02A} the size of the effect of the sample bias on the normalised expected values of the biased variables is associated with the relative size of the biased variable's coefficient, and;

H_{12A} there is no relationship between the size of the effect of sample bias on the biased variable and the relative size of the biased variable's coefficient.

As the plot showed a distinct relationship, non-linear regression was undertaken to quantify this relationship. The dependent variable was the percentage effect of sample bias on the normalised expected value coefficient of the biased variable. The independent variable was the normalised expected value of the biased coefficient with the equation being solved equal to:

$$\text{Percentage Bias} = a_0 (N E(\beta_b))^{b_1}$$

where : a_0 = intercept;

$N E(\beta_b)$ = the normalised expected value of the biased variable in the unbiased sampling distribution, and;

b_1 = the gradient.

The nonlinear regression used an iterative procedure where the best estimate of the coefficients a_0 and b_1 occurred when the relative change in the sum of squares of the residuals was less than or equal to $1.0E-8$. For each type and level of bias, a nonlinear model was fitted yielding six different estimates of a_0 and b_1 . Table 6 shows the estimated coefficients, standard errors, and fit of the respective models.

Table 6. Nonlinear regression: is the size of the coefficient a good predictor of the effect of sample bias?

Regression Coefficient and Model Fit	Type of Bias					
	Top Five Percent	Bottom Five Percent	Top Ten Percent	Bottom Ten Percent	Top Fifteen Percent	Bottom Fifteen Percent
a_0	-4.35	3.82	-6.86	6.91	-10.46	9.54
S.E. a_0	.30	.19	.41	.26	.39	.24
b_1	-1.00	-1.05	-1.04	-1.02	-1.00	-1.04
S.E. b_1	.03	.02	.03	.02	.02	.01
R^2	.99	.99	.99	.99	.99	.99

Model: Percentage Effect of Bias = $a_0 \times$ Normalised Unbiased Coefficient Expected Value b_1
 No estimates of b_1 were significantly different between models at the 95 percent level of confidence.

The results show that the size of the coefficient explains over 99 percent of the variation in the amount of sample bias. However, it should be noted that the model is only predicting nine data points with two parameters. Furthermore, none of the estimates of b_1 are significantly different at the 95 percent level of confidence, supporting the claim that there is a standard relationship between the independent and dependent variables (and the null hypothesis H_{02A}).

Consequently, these results suggest that although the earlier null hypothesis H_{02} was rejected at the 99 percent level of confidence, the sensitive nature of the MNL estimates to bias is predictable. The reason for this predictable sensitivity is not clear. However, it could be linked to the nonlinear relationship between the MNL choice probabilities and the parameter estimates.

The implication of this for marketers is that if they know the extent to which the sample they are fitting the MNL to is biased, then the coefficients can be transformed to remove this bias. However, the degree to which the nonlinear regression results of this study can be used to achieve this is unknown. They address fairly specific types of sample bias which may or may not be prevalent in practice. A number of other forms of sample bias (and for that matter sampling error

in general) and model forms need to be evaluated before a reliable procedure can be used.

5.3 Impact of Sample Bias upon the Parameter Estimates of the Unbiased Variables

Despite the finding that the MNL estimates of the biased variable appear to be sensitive to sample bias, the effect of the bias on the coefficients of the unbiased variables in the biased sampling distributions remains important. The MNL could still be modelling the error, but doing so for the coefficients of the unbiased variables. If this is the case, then we would observe a significant change in the normalised expected values of the coefficients of the unbiased variables relative to their value in the unbiased sampling distribution (alternative hypothesis H_{13}).

This would be unexpected given the mathematical analysis of omitted variable bias by Lee (1982) who found that omitting a relevant variable had no effect if there was no association between the omitted and included variables. The sample bias examined in this study is similar to the omitted variable bias analysed by Lee (1982), the difference being that the relevant variable is only partially omitted. Additionally, the method used here generates individuals with no interactions between the independent variables, so Lee's (1982) findings should also hold. Therefore, it can be viewed as an extension of Lee's (1982) research employing a different methodology and a greater number of independent variables in the model.

Consequently, the null hypothesis (H_{03}) states that there is no variation in the value of the unbiased coefficients in conditions of sample bias. This question is addressed by computing the average change in the unbiased coefficients caused by the sample bias on each of the independent variables. Table 7 shows the average change in the unbiased variables caused by sample bias on each of the nine independent variables.

The results indicate that none of the changes are significant at the 95 percent level of confidence meaning the null hypothesis (H_{03}) cannot be rejected. However, the direction of the change is consistently positive and progressively rises as the amount of bias increases. The reason for this is

not clear but could be due to the less than accurate normalisation of the expected values of the coefficients to account for scale parameter differences.

Table 7. How much does bias on one variable affect the value of the other variables?

Biased Variable	Top Five Percent	Bottom Five Percent	Type of Bias		Top Fifteen Percent	Bottom Fifteen Percent
			Top Ten Percent	Bottom Ten Percent		
1 Cost	+1.38 %	+1.37 %	+1.56 %	+1.92 %	+1.79 %	+2.72 %
2 Long	+0.99	+0.84	+1.55	+1.02	+1.94	+1.17
3 Todo	+0.58	+0.62	+0.84	+1.29	+0.95	+1.70
4 Climate	+0.80	+0.49	+1.07	+1.09	+1.11	+1.95
5 Food	+1.22	+1.19	+1.44	+2.13	+1.66	+3.04
6 Language	+1.13	+1.39	+1.68	+1.75	+2.00	+1.94
7 Scenery	+0.98	+0.68	+1.12	+0.87	+1.31	+1.45
8 Safety	+0.50	+0.44	+1.01	+0.60	+1.02	+1.90
9 Newtodo	+1.15	+0.49	+0.94	+0.58	+2.14	+0.57
Average	+0.97	+0.83	+1.25	+1.25	+1.55	+1.83

No differences significant at the 95 percent level of confidence.

Additionally, there does appear to be some inherent order to the degree of effect. Sample bias on variables one, five, and six consistently has a greater impact on the coefficients of the unbiased variables. Conversely, sample bias on variables three and eight consistently has the least effect. The possible cause of this cannot be identified by the standard deviation, kurtosis, or skewness of the sampling distribution of the biased variable or from the size of the coefficients. Nevertheless, despite the regular pattern being distinct and unexpected, it is not significant.

In summary, the results would suggest that the unbiased variables are not significantly affected by sample bias on another variable. This is consistent with expectations and again the null

hypothesis, that the MNL estimates of the unbiased coefficients are not sensitive to sample bias in another variable, cannot be rejected at the 95 percent level of confidence.

5.4 Sensitivity of the Unbiased Variables to Sampling Bias

The sensitivity of the unbiased coefficients to sample bias is investigated in this section. This contrasts with the previous section which investigated whether sample bias on some variables had a greater impact on the resulting coefficients than others. Ideally, the ratio between the expected values of the MNL estimates of the unbiased variables should not change in conditions of sample bias on one variable.

The previous section found that the amount of change in the unbiased variables was minimal (except possibly for scale parameter effects). Obviously, we would expect these findings to be evident in this section as well. Furthermore, the null hypothesis (H_{04}) suggests that the percentage changes experienced by each variable as a result of the sample bias should not be significantly different as this would mean that the ratio between the unbiased variables had not been maintained. In other words, if the effects are significantly different, then the effect of the sample bias would not be similar across the variables. In effect, such a finding would not only mean that the MNL model shifts the sampling error to the unbiased variables, but the extent to which this occurs depends upon factors other than the bias per se. This type of instability would not be desirable.

Table 8 shows the average change in the unbiased variables in the presence of sample bias. As expected, none of the differences between the effects of sample bias on the variables are significant at the 95 percent level of confidence. However, the average change in the coefficients of the unbiased variables progressively increases (albeit insignificantly). This indicates that even if the MNL's estimates of the unbiased variables reflect the level of bias, the ratio between the variables is maintained.

Again, the reason for the differences (caused by the proportional increase in bias from five to fifteen percent) evident in Table 8 are not clear. However, as previously suggested, they are

probably a result of changes in the value of the scale not accurately accounted for by the normalisation procedure.

Table 8. How sensitive is each variable to bias in other variables?

Unbiased Variable	<u>Type of Bias</u>					
	Top Five Percent	Bottom Five Percent	Top Ten Percent	Bottom Ten Percent	Top Fifteen Percent	Bottom Fifteen Percent
1 Cost	-0.21%	1.01%	-0.62%	-0.32%	-0.65%	-0.12%
2 Long	1.41	-0.62	1.85	-1.51	1.34	-3.24
3 Todo	0.17	-0.18	0.35	0.04	0.56	0.20
4 Climate	0.04	-0.09	-0.03	-0.18	0.01	-0.33
5 Food	0.07	0.06	0.06	0.07	-0.09	-0.05
6 Language	-0.03	0.90	0.43	2.18	1.18	3.79
7 Scenery	-0.49	0.14	-0.82	-0.42	-1.17	-1.07
8 Safety	0.23	-0.01	0.35	0.05	0.43	0.04
9 Newtodo	-0.77	-0.02	-1.05	-0.15	-1.42	0.12
Ave. Change	0.38	0.34	0.62	0.55	0.76	1.00

No significant differences at the 95 level of confidence.

Further to this, there also appears to be some degree of order to the amount of change experienced by each variable. Variables two and six are consistently the most affected by the bias while four, five, and eight are generally the least affected. The reasons for this are not clear but could again be a result of the incorrect normalisation. Nevertheless, these differences are not significant at the 95 percent level of confidence.

The results suggest that the null hypothesis (H_{05}) that all unbiased independent variables are equally affected by sample bias cannot be rejected at the 95 percent level of confidence.

5.5 Theoretical Standard Errors versus Standard Deviation of the Unbiased Sampling Distribution

A comparison of the theoretical standard errors of the coefficients of the original MNL model and the standard deviations of the unbiased sampling distribution was undertaken. Their likeness or otherwise serves as a guide of the accuracy of their reported values. The theoretical standard errors are those calculated by the software on the original data set of 164 individuals. These estimates are only asymptotically valid (Hensher and Johnson, 1981, p 48) when using ML estimation and assume a normal distribution of the parameter estimates. The asymptotic standard errors are computed by first, determining the second partial derivatives of the log likelihood with regard to the β 's, second, calculating the negative of the inverse of the matrix of these partial derivatives (this is the covariance matrix for the ML estimates), and finally, taking the square root of the diagonal components of this to yield the standard errors (Hensher and Johnson, 1981, p 48).

The standard errors of the coefficients in the unbiased sampling distribution are equal to their standard deviation. The method used here was specifically designed so that the calibration sample size of the original estimates of 164 was maintained. This meant these Monte Carlo standard deviations were directly comparable with the theoretical, asymptotic standard errors.

It would be desirable for these Monte Carlo standard deviations to approximate the asymptotic standard errors as they are theoretically identical. Nevertheless, if they do not, then one would have to question the use of asymptotic standard errors for smaller samples. The reported standard errors and the standard deviations of the unbiased sampling distributions are shown in table 9. Also included in this table are the measures of the normality of the Monte Carlo coefficients, namely, the kurtosis and skewness (the reason for which is explained below).

Table 9. Are the Asymptotic Standard Errors accurate?

Independent Variable	Asymptotic Standard Errors	Monte Carlo Standard Deviations	Difference	Kurtosis (Normal = 3)	Skewness (Normal = 0)
0 Intercept	0.061	0.085	+0.024	0.461	0.236
1 Cost	0.064	0.066	+0.002	0.015	-0.038
2 Long	0.071	0.136	+0.065	206.070	-8.938
3 Todo	0.068	0.086	+0.018	0.360	0.322
4 Climate	0.070	0.126	+0.056	287.662	11.568
5 Food	0.070	0.137	+0.067	195.983	8.597
6 Language	0.072	0.137	+0.065	195.671	-8.599
7 Scenery	0.072	0.128	+0.056	265.412	10.899
8 Safety	0.071	0.125	+0.054	281.339	11.367
9 Newtodo	0.067	0.089	+0.022	0.428	-0.192

Clearly, the software underestimates the size of the error in the coefficient estimates.

Nevertheless, it would appear that this variation is to a large extent caused by the non-normal nature of the distribution of the coefficients in the Monte Carlo unbiased sampling distribution as indicated by the kurtosis and skewness. Generally as the distribution of the coefficients nears normality, the discrepancy between the reported and actual standard errors decreases.

This finding, that the Monte Carlo generated sampling distributions are generally not normal, is unexpected given the fact that maximum likelihood estimators are asymptotically normal, "even if the distribution from which the observations were drawn is not normal" (Ramanathan, 1995, p 76). More disturbing is the fact that there is no clear explanation for some coefficient distributions (namely variable zero, one, three, and nine) being substantially more normal than others. Given the "dispersion-reducing tendencies of maximum likelihood estimation" for bootstrap estimates resulting in error in the estimates of the distribution (Efron, 1992, p 101), these Monte Carlo estimates of kurtosis and skewness could be expected to be exaggerated (that is if Efron's finding can be generalised to this method). This appears to occur for the sampling

distributions of the coefficients of variables two, four, five, six, seven, and eight, though again, no apparent explanation can be given for the other four coefficients having distributions less concentrated about the mean.

The effect of the scale parameter on the standard errors is not clearly explained in the modelling literature. Therefore, the standard errors reported above in table ttt have not been normalised as no obvious method for performing this is known. The results would suggest however, that the reported differences are more likely to be associated with the erroneous assumption of normality than scale differences. The fact that the differences are not linearly related would support this. Nevertheless, the truth or otherwise of this claim can only be known with certainty by controlling for scale parameter differences in future.

One of two conclusions can be made in the light of these results. First, the assumption of normality may be reasonable asymptotically, but unrealistic for the sample sizes commonly used in marketing. Coefficient normality is not evident here for samples of size 164. The question of the required sample size necessary to gain accurate estimates of the asymptotic standard errors seems to be important to solve.

Conversely, the Monte Carlo method employed in this study may be producing asymptotically non-normal sampling distributions. In other words, the Monte Carlo method may not be a viable technique for testing the accuracy of the standard errors. This assertion could be tested by generating Monte Carlo sampling distributions of coefficients where the MNL calibration sample size (in this case 164) is increased to a larger value, for example 10,000 or 100,000. If the Monte Carlo sampling distributions are still non-normal in these seemingly asymptotic conditions, then the Monte Carlo method could be deemed incapable of testing the accuracy of the standard errors. In either case, more needs to be known about the performance of the theoretical standard errors.

6.0 CONCLUSION

This study evaluated the sensitivity of the estimates produced by the MNL model to sample bias. The behaviour of the MNL in these conditions is important given the fact that the model is prevalently fitted to samples drawn from the population. A Monte Carlo approach was used to generate biased sampling distributions of the MNL coefficients. Sample bias was simulated by removing those individuals from the sample whose responses were most (or least) highly correlated with one of the independent variables. Therefore, this bias is analogous to partially omitting a relevant variable. An unbiased sampling distribution was also created as a benchmark.

Initially, the validity of the Monte Carlo method was considered. This was done by comparing the original or true coefficients with the expected value of the unbiased sampling distribution. As the Monte Carlo used the original coefficients to generate the unbiased sampling distribution, there should have been no significant differences between them (the null hypothesis H_{01}). The null hypothesis was confirmed at the 95 percent level of confidence subsequent to a linear adjustment of the values to account for the assumed scale differences.

Once the Monte Carlo method was corroborated, the main issues of interest were confronted. First, the effect of bias on the biased variable was investigated. The null hypothesis H_{02} stated that the percentage change in the value of the biased variable should be equal to the proportion of bias being simulated. The null hypothesis was unexpectedly rejected at the 99 percent level of confidence indicating the MNL coefficients are sensitive to bias. In some cases the sample bias changed the value of the biased coefficients by a massive 152 percent. Generally, the coefficients changed by between 50 and 100 percent more than the simulated bias. For example, when the sample was biased by five percent, the biased variable changed by ten percent. This is clearly not desirable.

Nevertheless, the biased variables moved in the expected direction. When the sample was biased

by removing those simulated individuals with a high correlation, the value of the coefficient declined in value and vice versa. Removing individuals with a low correlation with the biased variable consistently had a lesser effect on the coefficients than removing those with a high correlation. This difference could not be explained by the standard errors, the kurtosis, or the skewness of the sampling distributions of the biased coefficients.

However, there was an unexpected finding with regard to the sensitivity of the biased variables. A strong relationship was found between the percentage change in the normalised expected values of the biased variables and the relative size of their coefficients. Put simply, larger coefficients were less susceptible (although still sensitive) to sample bias than smaller coefficients. The reason for this relationship is not clear. In fact, the literature does not appear to recognise it. However, it could result from the nonlinear connection between the MNL predicted probabilities and the attributes of the alternatives, but this is speculation at this stage. An effect of this size could probably not be attributed to the imperfect scale parameter normalisation.

This is an important result as it suggests that although the MNL is sensitive to sample bias, the effect of the bias appears to be predictable and could be removed, or at the least minimised. More research could be undertaken to develop a method that adjusts the value of the parameters when the researcher recognises that the MNL calibration sample is biased.

The behaviour of the unbiased variables in response to the simulated sample bias was also examined. The average change in the value of the unbiased variables was not significant at the 95 percent level of confidence, supporting the claim that any change was independent of the bias (and confirming the null hypothesis H_{03}). Despite this, the direction and magnitude of the changes were consistent, although incorrect normalisation for the scale parameter could possibly account for this.

Additionally, the overall sensitivity of each variable to sample bias was examined. The results showed no variables were more prone to change as a consequence of sample bias. In the vast

majority of cases, the ratio between the coefficients of the unbiased variables was maintained supporting the null hypothesis H_{04} .

Although the expected values of the coefficients in the unbiased sampling distribution were close to their true or original values, the distributions of these estimates were not normal as expected, given the apparent asymptotic qualities of the maximum likelihood estimator. As a consequence, the asymptotic standard errors as reported in the software significantly underestimated the standard deviations of the variables in the unbiased sampling distribution. Further research may be required here to ascertain the reasons for the Monte Carlo not producing normally distributed coefficients. It could be due to the calibration sample size of 164 or the method used to generate individuals. As the standard errors play a prominent role in making statistical inferences about the coefficients, it is important that the minimum sample size required to yield reliable standard errors is determined.

Overall, the results suggest that the MNL is sensitive to sample bias. However, although they still fluctuate by more than expected, the larger effects are less sensitive to sample bias than the smaller ones. This is desirable and indicates that the larger effects are more reliable estimates of choice behaviour than smaller effects. It also means that reduction of error in the estimates would be better achieved by concentrating on gaining unbiased samples of the lesser effects.

Despite this, the effect of the bias is predictable being related to the size of the coefficient. This seems to convey the message that the sensitivity of the model to sample bias is more a reflection of its assumptions than of the variability in choice per se. The apparent independence of the estimation of the unbiased variables from the error supports this contention.

Furthermore, the model does appear to be measuring the effects it should. If the model was not reflecting the underlying choice behaviour of the sample, then it would produce estimates that varied randomly regardless of the size of the effect or the amount of bias. The predictable reaction of the biased variable's coefficients to sample bias and the independence of estimation of

the unbiased effects indicates that the model is producing estimates of effects where the ratio between them is maintained. This would suggest that when fitted to an unbiased sample, the model will produce estimates that reflect the effect being measured. The estimates of the smaller effects will also be more reliable in these circumstances than if gleaned from a biased data set.

7.0 REFERENCES

- Adamowicz, W., Louviere, J., & Williams, M. (1994). Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities. Journal of Environmental Economics & Management, 26, pp 271-292.
- Agresti, A. (1993). Distribution-Free Fitting of Logit Models with Random Effects for Repeated Categorical Responses. Statistics in Medicine, 12, 21, p 1969.
- Aguilera, A. & Gonzalez, V. (1993). A Multinomial Logit Analysis of Loan Targeting and Repayment at the Agricultural Development Bank of the Dominican Republic. Agricultural Finance Review, 53, p 55.
- Algorithm AS183. (1982). Applied Statistics, 31, pp 188-190.
- Amemiya, T. (1981). Qualitative Response Models: A Survey. Journal of Economic Literature, 19, pp 1483-1536.
- Asch, P., Malkiel, B.G., & Quandt, R. (1986). Market Efficiency in Racetrack Betting: Further Evidence and A Correction. Journal of Business, 59, 1, pp 157-160.
- Banerjee, B. (1982). Who Will Return and When? An Analysis of Rural-to-Urban Migrants in India. Oxford Bulletin of Economics & Statistics, 44, 4, pp 339-356.
- Batsell, R.R. & Polking, J.C. (1985). A New Class of Market Share Models. Marketing Science, 4, 3, pp 177-198.

- Bell, C.D., Roberts, R.K., & Park, W.M. (1994). A Logit Analysis of Participation in Tennessee's Forest Stewardship Program. Journal of Agricultural and Applied Economics, 26, 2, p 463.
- Ben-Akiva, M., Bradley, M., Morikawa, T., Benjamin, J., Novak, T., Oppewal, H., & Rao, V. (1994). Combining Revealed and Stated Preferences Data. Marketing Letters, 5, 4, pp 335-350.
- Ben-Akiva, M. & Lerman, S. (1991). Discrete Choice Analysis: Theory and Application to Travel Demand. Massachusetts: MIT Press.
- Berger, C.J., Olson, C.A., & Boudreau, J.W. (1983). Effects of Unions on Job Satisfaction: The Role of Work Related Values and Perceived Rewards. Organizational Behavior & Human Performance, 32, 3, pp 289-324.
- Berkson, J. (1953). A Statistically Precise and Relatively Simple Method of Estimating the Bio-Assay with Quantal Response Based on the Logistic Function. Journal of American Statistical Association, 48, pp 565-599.
- Billingsley, R.S. & Thompson, G.R. (1985). Determinants of Stock Repurchases by Bank Holding Companies. Journal of Bank Research, 16, 3, pp 128-135.
- Bishop, J. (1977). The Effect of Public Policies on the Demand for Higher Education. Journal of Human Resources, 12, 3, pp 285-307.
- Borjas, G. (1984). Race, Turnover, and Male Earnings. Industrial Relations, 23, 1, pp 73-89.

- Bunch, D.S. & Batsell, R.R. (1989) A Monte Carlo Comparison of Estimators for the Multinomial Logit Model. Journal of Marketing Research, 26, 1, pp 56-68.
- Bunn, D.N., Caudill, S.B., & Gropper, D.M. (1992). Crime in the Classroom: An Economic Analysis of Undergraduate Student Cheating Behaviour. The Journal of Economic Education, 23, 3, p 203.
- Cameron, T.A. (1985). A Nested Logit Model of Energy Conservation Activity by Owners of Existing Single Family Dwellings. Review of Economics & Statistics, 67, 2, pp 205-211.
- Cecchetti, S.G. (1986). The Frequency of Price Adjustment: A Study of the Newsstand Prices of Magazines. Journal of econometrics, 31, 3, pp 255-274.
- Chamberlain, G. (1980). Analysis of Covariance with Qualitative Data. Review of Economics Studies, XLVII, pp 225-238.
- Chambers, E.A. & Cox, D.R. (1967). Discrimination Between Alternative Binary Response Models. Biometrika, 54, pp 573-578.
- Chapman, R.G. & Staelin, R. (1982). Exploiting Rank Ordered Choice Set Data within the Stochastic Utility Model. Journal of marketing Research, 19, pp 288-301.
- Chen, Yu-Min & Anderson, E.E. (1993). Small Sample Properties of the Logit Maximum Likelihood Estimators: A Simulation Study. Journal of Applied Statistics, 20, 3, p 237.

- Chester, R., Kaufman, H.W., & Wolff, M.S. (1992). Use of Multiple Sensitivity Measurements and Logit Statistical Analysis to Assess the Effectiveness of A Potassium- Citrate- Contrainring Dentifrice in Reducing Dentinal Hypersensitivity. Journal of Clinical Periodontology, 19, 4, p 256.
- Chintagunta, Pradeep K. (1994). Heterogeneous Logit Model Implications for Brand Positioning. Journal of Marketing Research, 31, 2(May), p 304.
- Cooke, D., Craven, A.W., & Clarke, G.M. (1990). Basic Statistical Computing. (2nd Ed). London:Edward Arnold.
- Delton, A.H. & Marcis, J.G. (1986). The Determinants of Job Satisfaction for Young Male and Females. Atlantic Economic Journal, 14, 3, p 85.
- Dey, E.L. & Astin, A.W. (1993). Statistical Alternatives for Studying College Student Retention: A Comparative Analysis of Logit, Probit and Linear Regression. Research in Higher Education, 34, 5, p 569.
- Ederington, L.H. (1985). Classification Models and Bond Ratings. Financial Review, 20, 4, pp 237-262.
- Efron, B. (1992). Six Questions Raised by the Bootstrap. In: LePage, R. and Billard, L. (eds), Exploring the Limts of the Bootstrap. pp 99-126. Canada: John Wiley and Sons.
- Ehrenberg, R.G. & Marcus, A.J. (1982). Minimum Wages and Teenagers' Enrolment- Employment Outcomes: A Multinomial Logit Model. Journal of Human Resources, 17, 1, pp 39-58.

- Fader, P.S. (1993). Integrating the Dirichlet-Multinomial and Multinomial Logit Models of Brand Choice. Marketing Letters, 4, 2, p 99.
- Fader, P.S., Lattin, J. M., & Little, J.D.C. (1992). Estimating Nonlinear Parameters in the Multinomial Logit Model, Marketing Science, 11, 4, p 372.
- Falaris, E.M. (1982). Migration and Regional Wages. Southern Economic Journal, 48, 3, pp 670-686.
- Ferguson, W. & Yee, J. (1995). A Logit Model of Cotton Producer Participation in Professional Slout Programs. Journal of Sustainable Agriculture, 5, 3, p 87.
- Fields, G.S. & Mitchell, O.S. (1984). Qualitative Choice and the Blending of Discrete Alternatives. Review of Economics & Statistics, 66, 4, pp 547-555.
- Figlewski, S. (1979). Subjective Information and Market Efficiency in A Betting Market. Journal of Political Economy, 87, 1, pp 75-88.
- Finan, W.F. & Rappoport, P.N. (1982). The International Trade in Automobiles: Explaining Short-Run Changes in the Import Share of the United State Market, 1979-1981. Journal of Policy Modelling, 4, 1, pp 99-109.
- Fuller, W.C., Manski, C.F., & Wise, D.A. (1982). New Evidence on the Economic Determinants of Postsecondary Schooling Choices. Journal of Human Resources, 17, 4, pp 477-498.

- Garcia, C.V., Woodard, P.M., & Lee, B.S. (1995). A Logit Model for Predicting the Daily Occurrence of Human Caused Forest Fires. International Journal of Wildland Fire, 5, 2, p 101.
- Gensch, D.H. & Recker, W.W. (1979). The Multinomial, Multiattribute Logit Choice Model. Journal of Marketing Research, 16, pp 124-132.
- Gensch, D.H. & Soofi, E.S.(1995). Information-Theoretic Estimation of individual Consideration Set. International Journal of research in Marketing, 12, pp 25-38.
- Glaister, S. (1983). Some Characteristics of Rail Commuter Demand. Journal of Transport Economics & Policy, 17, 2, pp 115-132.
- Gonul, F. & Srinivasen, K. (1993). Modelling Multiple Sources of Heterogeneity in Multinomial Logit Models: Methodological and Managerial Issues. Marketing Science, 12, 3, p 213.
- Gordon, D.V., Lin, Z., Osberg, L., & Phipps, S. (1994). Predicting Probabilities: Inherent and Sampling Variability in the Estimation of Discrete-Choice Models. Oxford Bulletin of Economics & Statistics, 56, 1, pp 13-31.
- Gourieroux, C. & Monfort, A. (1993). Simulation-Based Inference: A Survey with Special Reference to Panel data Models. Journal of Econometrics, 59, pp 5-33.
- Green, P.E., Carmone, F.J., & Wachspress, D.P. (1977). On the Analysis of Qualitative Data in Marketing Research. Journal of Marketing Research, 14, pp 52-59.
- Greenspan, N.T. & Vogel, R.J. (1982). An Econometric Analysis of the Effects of Regulations in the Private Health Insurance Market. Journal of Risk & Insurance, 49, 1, pp 39-58.

- Grossman, J.B. (1984). The Occupational Attainment of Immigrant Women in Sweden. Scandinavian Journal of Economics, 86, 3, pp 337-351.
- Guadagni, P.M. & Little, J.D.C. (1983). A Logit Model of Brand Choice Calibrated on Scanner Data. Marketing Science, 2, 3, pp 203-238.
- Gunderson, M. (1980). Probit and Logit Model of Labour Force Participation. Industrial Relations, 19, 2, pp 216-220.
- Gupta, S. & Chintagunta, P. (1994). On Using Demographic Variables to Determine Segment Membership in Logit Mixture Models. Journal of Marketing Research, 31, 1(Feb), p 128.
- Hannan, T.H. (1983). Prices, Capacity, and the Entry Decision: A Conditional Logit Analysis. Southern Economic Journal, 50, 2, pp 539-550.
- Hasbrouck, J. (1985). The Characteristics of Takeover Targets: q and Other Measures. Journal of Banking & Finance, 9, 3, pp 351-362.
- Hausman, J. & McFadden, D. (1984). Specification Tests for the Multinomial Logit Model. Econometrica, 52, pp 1219-1240.
- Hayashi, Y., Isobe, T., & Tomita, Y. (1986). Modelling the Long-Term Effects of Transport and Land Use Policies on Industrial Locational Behaviour: A Discrete Choice Model System. Regional Science & Urban Economics, 16, 1, pp 123-143.
- Hensher, D.A. & Johnson, L.W. (1981). Applied Discrete-Choice Modelling. London: Croom Helm.

- Horowitz, J.L. (1981). Sampling Specification and Data Errors in Probabilistic Discrete-Choice Models, Appendix C. In: Hensher D.A. & Johnson L.W. (eds), Applied Discrete-Choice Modelling, pp 417-435. London: Croom Helm.
- Horowitz, J.L. & Louviere, J.J. (1995). What Is the Role of Consideration Sets in choice Modelling? International Journal of Research in Marketing, 12, pp 39-54.
- Hughes, G. & McCormick, B. (1982). Do Council Housing Policies Reduce Migration Between Regions? Economic Journal, 91, 364, pp 919-937.
- Iwarere, L.J. & Williams, J.E. (1991). A Micro-Market Analysis of Tenure Choice Using the Logit Model. The Journal of Real Estate Research, 6, 3, p 327.
- Jobber, D. (1994). Logit Model Analysis for Multivariate Categorical Data. Journal of Marketing Management, 10, 1/3, p 29.
- Johnson, E.J., Meyer, R.J., & Ghose, S. (1989). When Choice Models Fail: Compensatory Models in Negatively Correlated Environments. Journal of Marketing Research, 26, pp 255-270.
- Jones, J.M. & Landwehr, J. (1988). Removing Heterogeneity Bias from Logit Model Estimation. Marketing Science, 7, 1, pp 41-59.
- Jud, G.D. (1983). Real Estate Brokers and the Market for Residential Housing. Journal of the American Real Estate & Urban Economics, 11, 1, pp 69-82.
- Jung, T.Y. (1993). Ordered Logit Model for Residential Demand in Korea. Energy Economics, 15, 3, p 205.

- Kawninen, B. J. & Khawaja, M. S. (1995). Measuring Goodness of Fit for the Double-Bounded Logit Model. American Journal of Agricultural Economics, 77, 4, p 885.
- Kodde, D.A. (1986). Uncertainty and the Demand for Education. Review of Economics & Statistics, 68, 3, pp 460-467.
- Kerkvliet, J. (1994). Cheating By Economics Students: A Comparison of Survey Results. The Journal of Economics Education, 25, 2, p 121.
- Khraif, R.M. (1992). Permanent Versus Temporary Rural Migrants in Riyadh, Saudi Arabia: A Logit Analysis of Their Intentions of Future Mobility. Geojournal, 26, 3, pp 363.
- Killingsworth, M.R. & Reimers, C.W. (1983). Race, Ranking, Promotions, and Pay at a Federal Facility: A Logit Analysis. Industrial & Labor Relations Review, 37, 1, pp 92-107.
- Kling, C. L. & Herriges, J.A. (1995). An Empirical Investigation of the Consistency of Nested Logit Models with Utility Maximisation. American Journal of Agricultural Economics, 77, 4, p 875.
- Kodde, D.A. (1986). Uncertainty and the Demand for Education. Review of Economics & Statistics, 68, 3, pp 460-467.
- Kolchin, M.G. & Hyclak, T. (1984). Participation in Union Activities: A Multivariate Analysis. Journal of Labour Research, 5, 3, pp 255-262.
- Krieg, R.G. (1993). Black-White Regional Migration and the Impact of Education: A Multinomial Logit Analysis. The Analysis of Regional Science, 27, 3, p 211.

- Lawrence, E.C. & Arshadi, N. (1995). A Multinomial Logit Analysis of Problem Loan Resolution Choices in Banking. Journal of Money, Credit, and Banking, 27, 1, p 202.
- Lee, Lung-Fei (1982). Specification Error in Multinomial Logit Models: Analysis of the Omitted Variable Bias. Journal of Econometrics, 20, pp 197-209.
- Leppel, K. (1993). Logit Estimation of A Gravity Model of the College Enrolment Decision. Research in Higher Education, 34, 3, p 387.
- Levin, R.C. (1978). Allocation in Surface Freight Transportation: Does Rate Regulation Matter? Bell Journal of Economics, 9, 1, pp 18-45.
- Lewis, G.B. (1986). Gender and Promotion Chances of White Men and Women in Federal White-Collar Employment. Journal of Human Resources, 21, 3, pp 406-419.
- Linden, L., Rita, H., & Svojala, T. (1996). Logit Models for Estimating Lethal Temperatures in Apple. Hortscience, Feb, 31, 4, p 91.
- Lioukas, S.K. (1982). Choice of Travel Mode and the Value of Time in Greece: Some Robust Estimates. Journal of Transport Economics & Policy, 16, 2, pp 161-180.
- Lo, A.W. (1986). Logit Versus Discriminant Analysis: A Specification Test and Application to Corporate Bankruptcies. Journal of Econometrics, 31, 2, pp 151-178.
- Louviere, J. (1988). Analysing Decision Making: Metric Conjoint Analysis. Sage Publications.

Luce, R. (1959). Individual Choice Behaviour. New York: John Wiley.

Malhotra, N.K. (1984). The Use of Linear Logit Models in Marketing Research. Journal of Marketing Research, 21, pp 20-31.

Mandel, B., Gaudry, M., & Rothengatter, W. (1994). Linear or Non-Linear Utility Functions in Logit Models? The Impact on German High Speed Rail Demand Forecasts. Transportation research, Part B, 28, 2, p 91.

Manski, C.F. (1973). The Analysis of Qualitative Choice, Unpublished Ph.D. Thesis, Department of Economics, Massachusetts: MIT Cambridge.

Manski, C.F. & Thompson, T.S. (1986). Operational Characteristics of Maximum Score Estimation. Journal of Econometrics, 32, pp 85-108.

Marshall, G., Shroyer, L.W., & Grover, F.L. (1994). Bayesian-Logit Model for Risk Assessment in Coronary Artery Bypass Grafting. The Annals of Thoracic Surgery, 57, 6, p 1492.

Mauro, M.J. (1982). Strikes as a Result of Imperfect Information. Industrial & Labor Relations Review, 35, 4, pp 522-538.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behaviour. Frontiers in Econometrics, pp 105-142.

McFadden, D. & Reid, F. (1975). Aggregate Travel Demand Forecasting from Disaggregated Behavioural Models. Transportation Research Record, 534, pp 24-37.

- McFadden, D., Tye, W., & Train, K. (1977). An Application of Diagnostic Tests for the Independence from Irrelevant Alternatives Property of the Multinomial Logit Model. Transportation Research Record, 637, pp 39-45.
- McFadden, D. (1986). The Choice Theory Approach to Market Research. Marketing Science, 5, 4, pp 275-297.
- Morgan, J.B. (1986). A New Look at Debt Rescheduling Indicators and Models. Journal of International Business Studies, 17, 2, pp 37-54.
- Morrison, D.G. (1966). Testing Brand Switching Models. Journal of Marketing Research, 3, pp 401-409.
- Neelankawil, J.P., Mummalanerii, V., & Sessions, D.N. (1995). Use of Foreign Language and Models in Point Advertisements in East Asian Countries: A Logit Modelling Approach. European Journal of Marketing, 29, 4, p 24.
- Nijkamp, P. & Reggiuni, A. (1992). Impacts of Multiple Period Lags in Dynamic Logit Models. Geographical Analysis, 24, 2, p 159.
- Noreen, E.W. (1989). Computer-Intensive Methods for Testing Hypotheses: An Introduction. New York: John Wiley & Son.
- Perloff, J.M. (1981). The Duration of Unemployment in the Construction Industry. Journal of Labour Research, 2, 1, pp 111-131.
- Ramanathan, R. (1995). Introductory Econometrics with Applications. (3rd Ed). San Diego: The Dryden Press.

- Ripley, B.D. (1987). Stochastic Simulation. New York: John Wiley & Sons.
- Shields, M.P. & Tsui, S.W. (1983). The Probability of Another Child in Costa Rica. Economic Development & Cultural Change, 31, 4, pp 787-807.
- Small, K. & Hsiao, C. (1982). Multinomial Logit Specification Tests. Working Paper. Department of Economics, Princeton University.
- Smith, K.C., Savin, N.E., & Robertson, J.L. (1984). A Monte Carlo Comparison of Maximum Likelihood and Minimum Chi Square Sampling Distributions in Logit Analysis. Biometrics, 40, pp 471-482.
- Stewart, J.F., Harris, R.S., & Carleton, W.T. (1984). The Role of Market Structure in Merger Behaviour. Journal of Industrial Economics, 32, 3, pp 293-312.
- Swait, J. & Louviere, J.J. (1993). The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. Journal of Marketing Research.
- Thompson, G.L. (1977). Planning Considerations for Alternative Transit Route Structures. Journal of the American Institute of Planners, 43, 2, pp 158-168.
- Thurstone, L.L. (1927). A Law of Comparative Judgement. Psychological Review, 34, pp 278-286.
- Tzeng, G.H., Chen, J.J., & Lan, C.J. (1991). The Influence of Model Choice on Energy Conservation Strategies in Transport: Application of A Logit Model. Energy Economics, 13, 4, p 290.

Vanderporten, B. & Hall, W.C. (1983). Conditions Leading to Collective Negotiations Among Police. Journal of Labour Research, 4, 2, pp 137-148.

Waddell, P. (1992). A Multinomial Logit Model of Race and Urban Structure. Urban Geography, 13, 2, p 127.

Wittink, D.R. & Cattin, P. (1981). Alternative Estimation Methods for Conjoint Analysis: A Monte Carlo Study. Journal of Marketing Research, 18, pp 101-106.

APPENDIX A - FORTRAN 77 Programme

```

C*****
C*  BIASLOG.FOR    BIASLOG.FOR    BIASLOG.FOR    *
C*
C*  Programme designed to measure the effect of sample bias *
C*  on the estimates produced by the MNL choice model.      *
C*  Produces sampling distributions of MNL estimates under   *
C*  different conditions of sample bias.                    *
C*  Written in FORTRAN 77.                                  *
C*
C*                      BELL, FOX, and WRIGHT              *
C*                      1995                                *
C*
C*****

```

```

REAL RANRSP
REAL*8 COEFF(10),FSETS(24,12)
INTEGER AGGINT(12),IX,IY,IZ,TOTBIS,TOTDEL,BISVAR
CHARACTER*12 OUTFILE

```

```

C*****
C* The commands above set the characteristics of variables mentioned *
C* and set aside the appropriate amount of memory to store their *
C* values within the programme. *
C* RANRSP The random response to each choice set. *
C* COEFF(10) The 10 coefficient estimates from the NTELOGIT *
C* estimation program. Each time the random set of 164 *
C* individuals' aggregated responses to the 12 choice *
C* sets is produced, it is written into the design *
C* matrix. The NTELOGIT program then estimates the *
C* 10 coefficients (intercept plus nine independent vars) *
C* and these are transferred back to the main program *
C* through the matrix COEFF(10). *
C* FSETS(24,12) The matrix containing the aggregated random *
C* responses to the twelve choice sets and the design. *
C* Column 1 contains the choice set number, 2 the *
C* aggregated responses to the 12 choice sets, 3 the *
C* value of the intercept, 4-12 the presence (1) or *
C* absence (0) of the 9 independent variables. *
C* Odd rows represent the first alternative in each *
C* choice set. Even rows represent the second *
C* alternative in each choice set. *
C* AGGINT(12) The matrix of aggregated responses to the first *
C* alternative in the twelve choice sets. *
C* IX,IY,IZ The integer values of between 1 and 30000 required *
C* for the FUNCTION RANDOM. Before each sampling *
C* distribution is generated, these values should be *
C* changed to ensure that each set of distributions is *
C* not being produced from the same string of psuedo- *
C* random numbers. *
C* TOTBIS The total number of individuals to be created. This *
C* should include the base number of individuals to be *
C* produced (in this case 164), plus the number of *
C* individuals that will be deleted to simulate bias. *

```

```

c* TOTDEL      The total number of individuals to be removed in
c*              order to simulate bias.
c* BISVAR      The index of the variable on which the sample is to
c*              be biased.
c* OUTFILE     The name of the file to which the parameter
c*              estimates produced by NTELOGIT will be written.
c*****

```

PARAMETER (TOTBIS=182,TOTDEL=18,BISVAR=5,OUTFILE='V5BOT10.DAT')

```

c*****
c* THE FOLLOWING CHANGES CAN BE MADE TO TYPE OF BIAS SIMULATED BY
c* MAKING THE FOLLOWING CHANGES TO THE 'PARAMETER' COMMANDS
c* DIRECTLY BEFORE THIS MESSAGE.
c*
c* To set the level of bias:
c*   UNBIASED:          set TOTBIS equal to 164
c*                      set TOTDEL equal to 0
c*   FIVE percent bias: set TOTBIS equal to 173
c*                      set TOTDEL equal to 9
c*   TEN percent bias:  set TOTBIS equal to 182
c*                      set TOTDEL equal to 18
c*   FIFTEEN percent bias: set TOTBIS equal to 193
c*                      set TOTDEL equal to 29
c*
c* To bias different variables, change BISVAR to:
c*   1 - COST           (coef -.3128, std err .6066)
c*   2 - DISTANCE       (coef -.1306, std err .6426)
c*   3 - NUMBER         (coef .8809, std err .7061)
c*   4 - CLIMATE        (coef .8747, std err .6788)
c*   5 - FOOD           (coef .5992, std err .7001)
c*   6 - SPEAK          (coef .2692, std err .6996)
c*   7 - SCENERY       (coef .4293, std err .7192)
c*   8 - CRIME          (coef .9657, std err .7177)
c*   9 - VARIETY        (coef .4661, std err .7138)
c*
c* OUTFILE dictates the name of the file to which the coefficients
c* are stored.
c*****

```

INTEGER RESP(TOTBIS,12),DELREC(TOTBIS),DESIGN(12,9),DATUNT

```

c*****
c* More commands above which set the characteristics of variables
c* and set aside the appropriate amount of memory to store their
c* values within the programme.
c* RESP(TOTBIS,12) The matrix of random responses to the first
c*                  alternative of the 12 choice sets for the
c*                  TOTBIS individuals being generated.
c* DELREC(TOTBIS) The matrix that contains information used to
c*                determine the individuals to be removed. This
c*                information is the total number of times that
c*                each individual choose an alternative which
c*                contained the variable (BISVAR) upon which
c*                the sample was being biased.
c* DESIGN (12,9)   The design of the first alternative in the

```

```

c*           twelve choice sets. The nine values are the           *
c*           presence (1) or absence (0) of each independent       *
c*           variable for that alternative.                          *
c* DATUNT    The number of the unit to which data is read/       *
c*           written.                                             *
c*****

```

```

DATA DESIGN/0,0,1,1,0,1,1,1,0,0,0,1,1,0,1,0,0,1,0,
+0,0,1,1,1,1,1,0,1,0,0,1,0,0,0,1,1,0,0,0,1,0,1,1,0,1,1,
+1,1,0,1,0,0,1,0,0,0,0,1,0,0,0,1,0,1,1,0,1,1,1,0,0,1,1,1,0,1,0,
+0,1,0,0,1,1,1,0,0,0,1,0,1,1,0,1,1,0,0,0,1,1,1,0,1,0,0/

```

```

c*****
c* DESIGN:  The experimental design. Reads in the design matrix of *
c*          ONLY the first alternative in the twelve choice sets by *
c*          column then row. Only the design of the first alternative *
c*          is included because the second alternative is the exact   *
c*          opposite of the first.                                     *
c*****

```

```

IX=146
IY=535
IZ=21198

```

```

c*****
c* IX, IY, and IZ are the parameters of the random number          *
c* generator FUNCTION RANDOM(IX,IY,IZ). These must be set         *
c* to an integer between 1 and 30000.                              *
c*****

```

```

c***debugging*** write (6,*) 'Starting....'

```

```

DATUNT=11

DO 80 I=1,12
  II=2*I-1
  FSETS(II,1)=I
  FSETS(II+1,1)=I
  FSETS(II,3)=1.0
  FSETS(II+1,3)=0.0
  DO 81 J=1,9
    FSETS(II,J+3)= FLOAT(DESIGN(I,J))
    FSETS(II+1,J+3)=1.0-FLOAT(DESIGN(I,J))
  81 CONTINUE
80 CONTINUE

```

```

c*****
c* The above commands initialise the matrix FSETS(24,12) to make it *
c* compatible with the NTELOGIT estimation program.                 *
c* COLUMN 1 - the number of the choice set                          *
c* COLUMN 2 - this is left blank at this stage and once the       *
c* random individuals have been created and bias                   *
c* simulated, the aggregated responses to the 12                    *
c*

```

```

c*          choice sets (AGGINT(12)) are read into this          *
c*          column.                                             *
c*          COLUMN 3 - the intercept, 1 for 1st alternative (odd rows) *
c*          0 for 2nd alternative (even rows).                 *
c*          COLUMNS 4-12 - the presence (1) or absence (0) of the remaining *
c*          nine independent variables. Odd rows are equal to *
c*          the DESIGN(12,9) matrix, even rows are the inverse *
c*          of the preceding odd rows.                          *
c*****

```

```

OPEN(8,FILE=OUTFILE)
999 FORMAT(10f10.5)
20 FORMAT('+',i7)

```

```

c*****
c*          OPENS the OUTFILE and sets the FORMAT of the results. The value *
c*          of the estimated coefficients (COEFF(10)) is written to this file *
c*          later in this main programme.                          *
c*****

```

```

c***debugging*** write (6,*) 'Calling Generate....'

```

```

c*****
c*          The PRODUCT of n and nn sets the number of parameter estimates *
c*          returned by this programme. To help the user keep track of the *
c*          number of parameters produced, the value of n is printed on the *
c*          screen at each pass.                                    *
c*****

```

```

do 100 n=1,1000
do 110 nn=1,10

```

```

c*****
c*          CALLS the SUBROUTINE GENERATE which returns the random aggregated *
c*          responses to the first alternative for the twelve choice sets, AGGINT(12) *
c*****

```

```

CALL GENERATE(IX,IY,IZ,TOTBIS,TOTDEL,BISVAR,AGGINT,RESP,DELREC,
+ DESIGN)

```

```

c***debugging*** write out datafile
c***debugging*** open(11,file='VAC.DAT',STATUS='new')
c***debugging*** 121 format (I3,I4,10I3)

```

```

do 120 i=1,12

```

```

c***debugging*** write (11,121) I,AGGINT(I),1,(DESIGN(i,j),j=1,9)
c***debugging*** write (11,121) I,164-AGGINT(I),0,(1-DESIGN(i,j),j=1,9)

```

```

c*****

```

```

C*      Writes the randomly generated aggregated responses from the *
C*      SUBROUTINE GENERATE into the matrix FSETS. *
C*      FSETS contains the data and design matrices which are used *
C*      by the MNL estimation programme NTELOGIT in the *
C*      SUBROUTINE ESTIMATE. *
C*****

```

```

      II=2*I-1
      FSETS(II,2)=FLOAT(AGGINT(I))
      FSETS(II+1,2)=164.0-FLOAT(AGGINT(I))

```

```

120 continue

```

```

C*****
C*      CALL's the SUBROUTINE ESTIMATE which fits the MNL to *
C*      the randomly generated data and design (FSETS). Returns *
C*      the estimates of the 10 coefficients through COEFF. *
C*****

```

```

      CALL ESTIMATE(COEFF,FSETS)

```

```

c***debugging***   write (6,*) 'Returned from Generate....'

```

```

C*****
C*      The datafile is deleted when closed within the estimation routine. *
C*****

```

```

C*****
C*      WRITES the 10 COEFFS to the datafile, *
C*      OUTFILE. *
C*****

```

```

      WRITE(8,999) (COEFF(i),i=1,10)
110 continue

```

```

C*****
C*      WRITES n on screen so user can keep *
C*      track of number of estimates produced *
C*****

```

```

      write (6,20) n

```

```

100 continue

```

```

c***debugging***   write (6,*) 'Should end now ....'

```

```

      CLOSE (8,STATUS='KEEP')

```

```

      END

```

```

C*****
C*      THE END OF THE MAIN PROGRAMME *
C*      THE END OF THE MAIN PROGRAMME *

```

```

C*      THE END OF THE MAIN PROGRAMME      *
C*      THE END OF THE MAIN PROGRAMME      *
C*      THE END OF THE MAIN PROGRAMME      *
C*****

```

```

C*****
C*      SUBROUTINE GENERATE                  *
C*****

```

```

SUBROUTINE GENERATE(IX,IY,IZ,totbis,TOTDEL,bisvar,AGGINT,
+ resp,delrec,DESIGN)

```

```

C*****
C*      IX,IY and IZ are used by the random number generator      *
C*      TOTBIS is the total number of individuals generated        *
C*      TOTDEL is the number of individuals deleted                *
C*      BISVAR is the number of the variable used for biasing the sample *
C*      AGGINT is the aggregated output matrix of 12 numbers giving *
C*      the number of individuals (out of 164) that responded      *
C*      positively to the first alternative                          *
C*      RESP is the working area holding all responses from the full sample *
C*      DELREC is a working area to hold biasing information        *
C*      DESIGN contains the experimental design                      *
C*****

```

```

REAL RANRSP,CUTOFF(12)

```

```

INTEGER TOTBIS,BISVAR,TOTDEL
INTEGER RESP,AGGINT(12),DESIGN(12,9)
INTEGER FINRSP(164,12),DELREC(TOTBIS),IND
INTEGER DELCOU,IP,FIN,i

```

```

DIMENSION RESP(TOTBIS,12)

```

```

c***debugging*** OPEN(9,STATUS='NEW',FILE='DESIGN.DAT')
c***debugging*** DO 73 J=1,12
c***debugging*** WRITE(9,'(9I2)') (DESIGN(J,I),I=1,9)
c***debugging*** 73 CONTINUE
c***debugging*** CLOSE (9,STATUS='KEEP')

```

```

C*****
C* Parameter estimates gained from VAC.DAT used to yield probabilities *
C* for the two alternatives within each choice set. CUTOFF(1) is the *
C* is the probability that ALTERNATIVE ONE will be chosen from *
C* CHOICE SET ONE based on the original NTELOGIT parameter estimates. *
C*****

```

```

c***debugging*** write (6,*) 'Starting Generate....'

```

```

CUTOFF(1)=.66372888
CUTOFF(2)=.92773445
CUTOFF(3)=.17265963
CUTOFF(4)=.95538925
CUTOFF(5)=.04968568
CUTOFF(6)=.64114886
CUTOFF(7)=.50192799

```

```

CUTOFF(8)=.45641192
CUTOFF(9)=.41982691
CUTOFF(10)=.63221849
CUTOFF(11)=.94095225
CUTOFF(12)=.08343799

```

```

C*****
c* Sets the matrix RESP(i, j). *
C*****

```

```

      DO 30 i=1,totbis
        DO 40 j=1,12
          RESP(i,j) = 0
40    CONTINUE
30    CONTINUE

```

```

C*****
c* Sets the matrix DELREC(i). *
C*****

```

```

      DO 20 i=1,totbis
        DELREC(i) = 0
20    CONTINUE

```

```

c***debugging***   write (6,*) 'Initialization complete..'

```

```

C*****
c* Uses the random number function to generate an array of TOTBIS x12 *
c* random 0/1 responses. Changes the original uniform random numbers *
c* into a discrete format using CUTOFF's. *
C*****

```

```

      DO 100 i=1,totbis
        DO 50 j=1,12
          RANRSP = RANDOM(IX,IY,IZ)
          IF(RANRSP.LE.CUTOFF(j)) THEN
            RESP(i,j)=1
          ELSE
            RESP(i,j)=0
          ENDIF
50    CONTINUE
100   CONTINUE

```

```

c***debugging***   write (6,*) 'Random array generated....'

```

```

C*****
c* Calculates the number of times that each simulated *
c* individual selects an alternative containing the independent *
c* variable upon which the sample is being biased. This is the *
c* same as computing the correlation between the bias variable *
c* and each individuals responses to the twelve choice sets. *
c* Each time the individual (RESP) selects an alternative in *
c* the DESIGN containing the bias variable (BISVAR), one is *
c* added to DELREC. DELREC records for each individual the *
c* number of times the reponse and design correspond. *
c* DELREC ranges from 0 representing a simulated individual *

```

```

c* who did not select one alternative containing BISVAR, to 12 *
c* representing an individual who selected all alternatives *
c* containing BISVAR. IND equals the count (from 1 to TOTBIS) *
c* of the number of the individual in the matrix RESP. *
c*****

```

```

IND=0

```

```

do 291 ind=1,totbis
  DO 290 i=1,12
    IF(RESP(ind,i).EQ.DESIGN(i,bisvar)) THEN
      DELREC(IND)=DELREC(IND)+1
    ENDIF
  290 continue
291 CONTINUE

```

```

c***debugging*** write (6,*) 'Successfully past first GOTO loop...!'

```

```

c*****
c* Simulates sample bias by removing those individuals whose *
c* responses are highly or lowly correlated with the bias *
c* variable (BISVAR). The programme searches through and *
c* marks those individuals in the array DELREC(IND) whose *
c* record equalled LIM until the number required to be *
c* removed (when DELCOU equals TOTDEL) had been *
c* reached. *
c* *
c* IND = number of individuals in array DELREC from *
c* 1 to TOTBIS. *
c* DELREC = for each IND, the number of times an alternative *
c* containing the BISVAR had been selected. *
c* LIM = sets the type of bias to be simulated. *
c* LIM is set to 12 when individuals are to be *
c* removed from the top of the sample distribution, *
c* 0 for the bottom. *
c* DELCOU = counts the number of individuals in the array *
c* DELREC(IND) whose record equalled LIM. *
c*****

```

```

300 IND=0
    DELCOU=0
    LIM=0

```

```

320 IND=IND+1

```

```

c*****
c* Marks each individual (IND) whose DELREC equalled LIM *
c* with and -1. *
c*****

```

```

IF(DELREC(IND).EQ.LIM)THEN
  DELCOU=DELCOU+1
  DELREC(IND)=-1
ENDIF

```

```

c*****
c* Stops marking individuals for deletion when the number *
c* required to be deleted (TOTDEL) is reached (DELCOU). *
c*****

```

```

340 IF(DELCOU.EQ.TOTDEL)THEN
      GOTO 360
ENDIF

```

```

C*****
C*  Once the array DELREC(IND) has been searched but the  *
C*  number required to be deleted has not been reached, LIM is  *
C*  either reduced by 1 (for top bias) or increased by 1 (for  *
C*  bottom bias). The array DELREC(IND) is then searched  *
C*  again and individuals marked until TOTBIS is reached.  *
C*****

```

```

      IF(IND.EQ.totbis)THEN
        LIM=LIM+1
        IND=0
      ENDIF

```

```

350 GOTO 320

```

```

360 CONTINUE

```

```

C***debugging***   write (6,*) 'Successfully past second GOTO loop....'

```

```

C*****
C*  Copies RESP() to FINRSP() missing out those individuals  *
C*  who had been marked for deletion in the matrix DELREC(). *
C*****

```

```

      ind=1
      do 700 i=1,totbis
        if(delrec(i).ge.0) then
          do 710 j=1,12
            FINRSP(ind,j)=RESP(i,j)
          710 continue
          ind=ind+1
        endif
      700 continue

```

```

C*****
C*  Aggregates the results of the array FINRSP(164,12).  *
C*****

```

```

      DO 810 j=1,12
        AGGINT(j) = 0
      810 CONTINUE

```

```

      DO 950 i=1,164
        DO 940 j=1,12
          AGGINT(j) = AGGINT(j) + finrsp(i,j)
        940 CONTINUE
      950 CONTINUE

```

```

C***debugging***   OPEN(8,FILE='AGG')
C***debugging***   WRITE(8,999) (AGGINT(j),j=1,12)
C***debugging***   999 FORMAT(I4)
C***debugging***   CLOSE (8,STATUS='KEEP')

```

RETURN

END

```
C*****  
C*          FUNCTION RANDOM          *  
C*****
```

FUNCTION RANDOM(IX,IY,IZ)

```
C*****  
C* Algorithm AS183 from Applied Statistics 31: 188-190 (1982). Returns *  
C* a psuedo-random number between 0 and 1. The parameters IX, IY, and *  
C* IZ must be set to integer values between 1 and 30000 before the *  
C* first entry. Integer arithmetic up to 30323 is needed. The dummy *  
C* parameter is not used. *  
C*****
```

```
IX=171*MOD(IX,177)-2*(IX/177)  
IY=172*MOD(IY,176)-35*(IY/176)  
IZ=170*MOD(IZ,178)-63*(IZ/178)  
IF (IX.LT.0) IX=IX+30269  
IF (IY.LT.0) IY=IY+30307  
IF (IZ.LT.0) IZ=IZ+30323  
RANDOM=AMOD(FLOAT(IX)/30269.0+FLOAT(IY)/30307.0+  
+ FLOAT(IZ)/30323.0,1.0)
```

RETURN

END

```
C*****  
C*          SUBROUTINE ESTIMATE          *  
C*****
```

SUBROUTINE ESTIMATE(COEFF,FSETS)

REAL*8 COEFF(10)

```
C*****  
C          NTELOGIT MNL ESTIMATION PROGRAMME  
C          MULTINOMIAL LOGIT BY ITERATIVELY REWEIGHTED LEAST SQUARES  
C  
C          VERSION 4.0 GEORGE G. WOODWORTH 2/07/88  
C          COMMAND INPUT BY CAROL GILBERT  
C          Extended Format to 512 Characters - ber 89060  
C          Add DELETE - ber 89061  
C          Add BETSET Michael Fox 25 Aug 89  
C          Replace END with EXECUTE  
C          and STOP Michael Fox 14 Nov 89  
C  
C          DIMENSIONS ARE AT LEAST:  
C          ZPZ(NVARS*(NVARS-1)/2),BETA(NVARS-2),XPB(ALTMAX),PHAT(ALTMAX)  
C          ZTOT(NVARS-1),Z(ALTMAX,NVARS),ALTPTR(ALTMAX),  
C          VARPTR(NVARS),SWEPT(NVARS-1),VARNAM(NVARS),  
C          NZPTR(NVARS)  
C
```

```

C*****
PARAMETER (NVARMAX=13, NALTMAX=5)
REAL*8 ZPZ(NVARMAX*(NVARMAX-1)/2),Z(NALTMAX,NVARMAX),
+ BETA(NVARMAX-2),ZTOT(NVARMAX-1),XPB(NALTMAX),
+ PHAT(NALTMAX),TOLEPS,CVGEPS,LRCHIS,LRCH1,
+ RBUF(NVARMAX),FSETS(24,12)
INTEGER
VARPTR(NVARMAX),SWEPT(NVARMAX-1),ALTPTR(NALTMAX),CMDUNT,
+ DATUNT,ALTMAX,NVARS,JGRP,JFRQ,DFERR,NZPTR(NVARMAX),
+ PAGENO,LOGUNT,LSTUNT,PAGSIZ,ROWPAG,COLPAG,NDEL,
+ DMASK(NVARMAX),NSETS
CHARACTER*8 VARNAM(NVARMAX),DELNAM(NVARMAX),RJVNAM(NVARMAX)
CHARACTER*64 CMDFIL
CHARACTER*127 ERRMSG
CHARACTER*512 DATFMT
LOGICAL ERROR,CONVRG,COVMAT,COVR4T,POISON
DATA CMDUNT,DATUNT,LOGUNT,LSTUNT/10,11,6,6/,
# TOLEPS,CVGEPS,MAXITR/1.D-10,1.D-6,20/,
# PAGSIZ,ROWPAG,COLPAG/58,53,6/

CMDFIL='VAC.CMD'
c OPEN (CMDUNT,FILE=CMDFIL,STATUS='OLD')

200 ITER = 0
LOGUNT = 6
LSTUNT = 6
ERROR = .FALSE.
ALTMAX = NALTMAX
DO 300 K=1, NVARMAX-1
SWEPT(K) = 1
300 CONTINUE
NVARS = 0
NDEL=0
DO 310 K=1, NVARMAX-2
BETA(K)=0.0D0
310 CONTINUE

CALL CMDINP(CMDFIL,CMDUNT,DATUNT,DATFMT,NVARMAX,
# NVARS,VARNAM,VARPTR,BETA,ERROR,ERRMSG,
# LOGUNT,LSTUNT,PAGSIZ,ROWPAG,COLPAG,COVMAT,
# COVR4T,POISON,NDEL,DELNAM,DMASK)
IF (ERROR) GO TO 900

c write (6,*) ' Commands input ...'

100 CALL WCSSP(DATUNT,DATFMT,VARPTR,ALTMAX,NVARS,RBUF,BETA,ZPZ,
# LRCHIS,DFERR,Z,ZTOT,XPB,PHAT,ALTPTR,POISON,ERROR,
# ERRMSG,NZPTR,NDEL,DMASK,LRCH1,NSETS,FSETS)
IF (ERROR) GO TO 900

c write (6,*) ' WCSSP returned successfully.'

CALL UPDATE(ZPZ,BETA,NVARS,VARPTR,TOLEPS,CVGEPS,MAXITR,
# ITER,CONVRG,SWEPT,ERROR,ERRMSG)
IF (ERROR) GO TO 900

c write (6,*) ' UPDATE returned successfully.'

IF (.NOT. CONVRG) GO TO 100

```



```
ERROR=.FALSE.
TEST=.TRUE.
LSTPTR=.FALSE.
LOGPTR=.FALSE.
COVMAT=.FALSE.
COVR4T=.FALSE.
POISON=.FALSE.
GRPSET=.FALSE.
FRQSET=.FALSE.
BETSET=.FALSE.
NODATA=.TRUE.
NV=0
ND=0
IO=0
```

```
10 NV=12
```

```
20 VARNAM(1)='GROUP '
   VARNAM(2)='FREQ '
   VARNAM(3)='INTCEPT '
   VARNAM(4)='COST '
   VARNAM(5)='LONG '
   VARNAM(6)='TODO '
   VARNAM(7)='CLIMATE '
   VARNAM(8)='FOOD '
   VARNAM(9)='LANGUAGE'
   VARNAM(10)='SCENERY '
   VARNAM(11)='SAFETY '
   VARNAM(12)='NEWTODO '
```

```
30 FORMAT=(F3.0,F4.0,10F3.0)
```

```
40 DATFIL='VAC.DAT'
```

```
C
```

```
C Don't need to open this, as it is still open after writing it.
```

```
C OPEN (DATUNT,FILE=DATFIL,STATUS='OLD')
```

```
   NODATA=.FALSE.
```

```
C REWIND(DATUNT)
```

```
50 GRPVAR='GROUP '
   grpset=.true.
   JGRP=1
```

```
60 FRQVAR='FREQ '
   frqset=.true.
   JFRQ=2
```

```
* STOP **** Added 14 Nov. 1989 MFF
```

```
*
```

```
250 CONTINUE
```

```
c ERROR = .TRUE.
```

```
c ERRMSG = ' End of Command File reached '
```

```
   do 70 i=1,12
     VARPTR(i)=i
70 continue
```

```
RETURN
```

```
END
```

```

C*****
C WCSSP (WEIGHTED, CENTERED SUMS OF SQUARES AND PRODUCTS) *
C INPUTS: *
C DATUNT = UNIT NUMBER OF DATA FILE *
C DATFMT = DATA FORMAT *
C VARPTR = VECTOR OF POINTERS TO GROUP,FREQ,X(1),...,X(BETDIM) *
C ALTMAX = MAXIMUM NUMBER OF ALTERNATIVES IN ANY GROUP *
C NVAR = NUMBER OF VARIABLES (INCLUDING GROUP AND FREQ) *
C BETA = CURRENT PARAMETER VECTOR *
C OUTPUTS: *
C ZPZ = WEIGHTED Z'Z IN TRIANGULAR STORAGE MODE *
C DFERR = DEGREES OF FREEDOM FOR ERROR *
C NSETS = NUMBER OF CHOICE SETS IN DATA FILE *
C WORK ARRAYS: *
C Z = WORK AREA FOR DATA SUBMATRIX *
C ZTOT = WORK AREA FOR RUNNING TOTALS *
C XPB = WORK AREA FOR X'BETA *
C PHAT = WORK AREA FOR FITTED PROBABILITIES *
C ALTPTR = WORK AREA FOR ALTERNATIVE POINTERS *
C NZPTR = WORK AREA FOR NONZERO ELEMENT POINTERS *
C ERROR REPORTS: *
C ERROR = ERROR FLAG (LOGICAL) *
C ERRMSG = ERROR MESSAGE ASSOCIATED WITH ERROR FLAG *
C*****

```

```

SUBROUTINE WCSSP(DATUNT,DATFMT,VARPTR,ALTMAX,NVAR,RBUF,BETA,
# ZPZ,LRCHIS,DFERR,Z,ZTOT,XPB,PHAT,ALTPTR,POISON,
# ERROR,ERRMSG,NZPTR,NDEL,DMASK,LRCH1,NSETS,FSETS)
LOGICAL EOD,ERROR,POISON
INTEGER NVAR,VARPTR(NVAR),ALTMAX,DATUNT,DFERR,ALTPTR(ALTMAX),
# ALTERN,BETDIM,IZOLD,IZ,IZPZ,IALT,JZ,JGRP,JFRQ,
# JC,JR,JB,JCOL,JROW,JCBASE,NALT,NZPTR(NVAR),NNZ,
# NDEL,DMASK(NVAR),NREAD,NSETS,LINE
REAL*4 GROUP,LSTGRP
REAL*8 ZPZ(NVAR*(NVAR-1)/2),Z(ALTMAX,NVAR),BETA(NVAR-2),
# ZTOT(NVAR-1),XPB(ALTMAX),LRCHIS,LRCH1,RFACT,FSETS(24,12),
# PHAT(ALTMAX),PTOT,FTOT,FHAT,ZROW,ZCOL,WTOT,RBUF(NVAR)
CHARACTER*512 DATFMT
CHARACTER*127 ERRMSG

```

```

* INITIALIZE
C REWIND (DATUNT)
EOD = .FALSE.
GROUP = -9999.
ALTERN = 0
IZOLD = 0
BETDIM = NVAR - 2
JGRP = VARPTR(1)
JFRQ = VARPTR(2)
DFERR = -BETDIM
LRCHIS = 0.D00
LRCH1 = 0.0D00
NREAD = NVAR + NDEL
NSETS = 0
LINE=0

```

```

DO 90 IZPZ = 1,NVAR*(NVAR-1)/2
ZPZ(IZPZ) = 0.D00
90 CONTINUE

```

```

* FILL DATA BUFFER UNTIL GROUP NUMBER CHANGES
100 LSTGRP = GROUP
    IZ = MOD(IZOLD,ALTMAX)+1
C*****
C Add the capability to delete unwanted variables from input stream
C - BER 89.06.04
C*****

C
C Read the entire dataline, delete the unnecessary data
C
C IF (NDEL .GT. 0) THEN
C READ (11,DATFMT,END=110,ERR=110) (RBUF(J),J=1,NREAD)
C DO 105 J=1,NVARS
C Z(IZ,J) = RBUF(DMASK(J))
C 105 CONTINUE
C ELSE
C
C Use all the data, DELETE option NOT specified, avoid unnecc. proce
C
C WRITE(*,*) 'Before read ... '
C LINE=LINE+1
C DO 106 J=1,12
C Z(IZ,J)=FSETS(LINE,J)
106 CONTINUE
C IF(LINE.EQ.24) GOTO 110

C READ (11,DATFMT,END=110,ERR=110) (Z(IZ,J),J=1,NVARS)
C WRITE(*,*) 'After read, IZ = ',IZ
C ENDIF

C IF NOT AT END OF DATA FILE THEN
C GROUP = Z(IZ,JGRP)
C IF (MOD(INT(GROUP),10000).EQ.0) WRITE(6,*) 'Processing ',GROUP
C WRITE (6,*) 'Processing ',GROUP
C IF (LSTGRP .EQ. GROUP) THEN
C IZOLD = IZ
C ALTERN = ALTERN+1
C IF (ALTERN .GT. ALTMAX) THEN
C ERROR = .TRUE.
C ERRMSG = ' Problem in DATAFILE: Too many alternatives'
C RETURN
C END IF
C ALTPTR(ALTERN) = IZ
C GO TO 100
C END IF
C GOTO 120
C ELSE (AT END OF DATA FILE)
110 EOD = .TRUE.
C END IF

*-----
* INCREMENT Z'Z MATRIX
120 IZOLD = IZ
NALT = ALTERN
IF (NALT .GT. 0) THEN
NSETS = NSETS + 1
DFERR = DFERR + NALT - 1
PTOT = 0.D00

```

```

FTOT = 0.
WTOT = 0.D00

DO 125 JZ = 1, NVAR
  ZTOT(JZ) = 0.D00
125 CONTINUE

DO 140 IALT = 1, NALT
  IZ = ALTPTR(IALT)
  XPB(IALT) = 0.D00
  JB = 0
  DO 130 JB = 1, BETDIM
    JZ = VARPTR(JB+2)
    XPB(IALT) = XPB(IALT) + Z(IZ, JZ)*BETA(JB)
130 CONTINUE
  PHAT(IALT) = DEXP(XPB(IALT))
  PTOT = PTOT + PHAT(IALT)
  FTOT = FTOT + Z(IZ, JFRQ)
140 CONTINUE

LRCH1 = LRCH1 + RFACT(FTOT, POISON) - FTOT*DLOG(FLOAT(NALT))
LRCHIS = LRCHIS + RFACT(FTOT, POISON)
DO 190 IALT = 1, NALT
  IZ = ALTPTR(IALT)
  PHAT(IALT) = PHAT(IALT)/PTOT
  FHAT = FTOT*PHAT(IALT)
  LRCHIS = LRCHIS - RFACT(Z(IZ, JFRQ), POISON)
+       + Z(IZ, JFRQ)*DLOG(PHAT(IALT))
  LRCH1 = LRCH1 - RFACT(Z(IZ, JFRQ), POISON)
  Z(IZ, JFRQ) = XPB(IALT) + Z(IZ, JFRQ)/FHAT - 1.
  WTOT = WTOT + FHAT

JCBASE = -1
NNZ = 0
DO 180 JC = 2, NVAR
  JCOL = VARPTR(JC)
  JCBASE = JCBASE + JC - 2
  ZCOL = Z(IZ, JCOL)*FHAT
  IF (ZCOL .NE. 0) THEN
    NNZ = NNZ + 1
    NZPTR(NNZ) = JC
    ZTOT(JCOL) = ZTOT(JCOL) + ZCOL
    DO 170 JNZ = 1, NNZ
      JR = NZPTR(JNZ)
      JROW = VARPTR(JR)
      IZPZ = JR + JCBASE
      ZPZ(IZPZ) = ZPZ(IZPZ) + Z(IZ, JROW)*ZCOL
170 CONTINUE
    END IF
180 CONTINUE
190 CONTINUE

JCBASE = -1
DO 210 JC = 2, NVAR
  JCOL = VARPTR(JC)
  JCBASE = JCBASE + JC - 2
  ZCOL = ZTOT(JCOL)/WTOT
DO 200 JR = 2, JC
  JROW = VARPTR(JR)
  IZPZ = JR + JCBASE

```

```

        ZPZ(IZPZ) = ZPZ(IZPZ) - ZTOT(JROW)*ZCOL
200  CONTINUE
210  CONTINUE

```

```

END IF

```

```

*-----

```

```

IF (.NOT. EOD) THEN
  ALTERN = 1
  ALTPTR(ALTERN) = IZOLD
  GO TO 100
ELSE
  RETURN
END IF
END

```

```

C*****

```

```

C UPDATE
C UPDATES THE BETA VECTOR.
C INPUTS:
C   ZPZ      (MODIFIED)
C   BETA     (MODIFIED)
C   NVAR     (MODIFIED)
C   VARPTR
C   TOLEPS, CVGEPS TOLERANCE AND CONVERGENCE EPSILONS
C   MAXITR   MAXIMUM NUMBER OF ITERATIONS
C   ITER     (MODIFIED)
C OUTPUTS:
C   CONVRG   .TRUE. WHEN BETA CONVERGES
C   ERROR, ERRMSG

```

```

C*****

```

```

SUBROUTINE UPDATE (ZPZ,BETA,NVAR,VARPTR,TOLEPS,CVGEPS,
#                MAXITR,ITER,CONVRG,SWEPT,ERROR,ERRMSG)
INTEGER NVAR,VARPTR(NVAR),MAXITR,ITER,SWEPT(NVAR-1),JPIVOT
REAL*8 ZPZ(NVAR*(NVAR-1)/2),BETA(NVAR-2),TOLEPS,CVGEPS,
#DELTA,MODDLT
LOGICAL CONVRG,ERROR
CHARACTER*127 ERRMSG

```

```

CONVRG = .FALSE.
ITER = ITER + 1

```

```

DO 100 JPIVOT = 2,NVAR-1
  CALL SWEEP(ZPZ,NVAR-1,JPIVOT,SWEPT,TOLEPS,ERROR,ERRMSG)
  IF (ERROR) RETURN
100 CONTINUE

```

```

c write (6,*) ' UPDATE: Step 1.....'

```

```

C TEST FOR CONVERGENCE
MODDLT = 0.D00
DO 110 JB = 1, NVAR-2
  DELTA = BETA(JB)
  JZPZB = JB*(JB+1)/2+1
  JZPZV = JZPZB + JB
  BETA(JB) = -ZPZ(JZPZB)
  DELTA = (DELTA-BETA(JB))**2/ZPZ(JZPZV)
  MODDLT = MODDLT + DELTA
110 CONTINUE
MODDLT = DSQRT(MODDLT/(NVAR-2))

```

```

c write (6,*) ' UPDATE: Step 2.....'

```

```

IF (MODDLT .LE. CVGEPS) THEN
  CONVRG = .TRUE.
END IF

```

```

IF (ITER .GT. MAXITR) THEN
  CONVRG = .TRUE.
  ERROR = .TRUE.
  ERRMSG = 'MAXIMUM ITERATIONS EXCEEDED.'
END IF

```

```

RETURN
END

```

```

SUBROUTINE SWEEP(Z,IORDER,IP,SWEPT,EPS,ERROR,ERRMSG)

```

```

C*****
C REVERSIBLE UPPER TRIANGULAR SWEEP FROM J.H. GOODNIGHT, *
C THE SWEEP OPERATOR: ITS IMPORTANCE IN STATISTICAL COMPUTING *
C
C GEORGE WOODWORTH 6/3/85 *
C*****

```

```

DOUBLE PRECISION Z(1),EPS,B,C,D
INTEGER SWEPT(1)
LOGICAL ERROR
CHARACTER*127 ERRMSG
IZP = IZADR(IP,IP)
D = Z(IZP)
IF (D .LT. EPS) THEN
  ERROR = .TRUE.
  ERRMSG = 'DESIGN MATRIX IS SINGULAR.'
RETURN
ENDIF

```

```

DO 100 IR = 1, IORDER
  IF (IR .NE. IP) THEN
    IZB=IZADR(IR,IP)
    B = Z(IZB)/D
    IF (IR .GT. IP) THEN
      B = SWEPT(IR)*SWEPT(IP)*B
    ENDIF
    DO 90 IC = IR, IORDER
      IF (IC .NE. IP) THEN
        IZC = IZADR(IC,IP)
        C = Z(IZC)
        IF (IC .LT. IP) THEN
          C = SWEPT(IC)*SWEPT(IP)*C
        ENDIF
        IZ = IZADR(IR,IC)
        Z(IZ) = Z(IZ) - B*C
      ENDIF
    90 CONTINUE
  ENDIF
100 CONTINUE

```

```

DO 200 IR = 1, IORDER
  IF (IR .NE. IP) THEN
    IZ = IZADR(IR,IP)
    Z(IZ) = Z(IZ)/D
    IF (IR .LT. IP) THEN

```

```

      Z(IZ) = - Z(IZ)
    ENDIF
  ENDIF
200 CONTINUE

```

```

      Z(IZP) = 1/D
      SWEPT(IP) = -SWEPT(IP)

```

```

      RETURN
    END

```

```

      FUNCTION IZADR(I1,I2)

```

```

C-----
C RETURNS ROW-MAJOR UPPER TRIANGULAR ADDRESS FOR ROW I1, COL I2.
C
C GEORGE WOODWORTH 6/3/85
C-----

```

```

      IR=MIN0(I1,I2)
      IC=MAX0(I1,I2)
      IZADR=IR+(IC*(IC-1))/2
      RETURN
    END

```

```

      SUBROUTINE UPCASE(LEN,LETTER)

```

```

      INTEGER LEN
      CHARACTER*1 UCASE(26),LCASE(26)
      CHARACTER*1 LETTER(LEN)

```

```

      DATA UCASE/'A','B','C','D','E','F','G','H','I','J','K','L','M'
+      , 'N','O','P','Q','R','S','T','U','V','W','X','Y','Z'/
      DATA LCASE/'a','b','c','d','e','f','g','h','i','j','k','l','m'
+      , 'n','o','p','q','r','s','t','u','v','w','x','y','z'/

```

```

      DO 100 N=1,LEN
        DO 110 J=1,26
          IF(LETTER(N).EQ.LCASE(J)) LETTER(N)=UCASE(J)
110 CONTINUE
100 CONTINUE
      RETURN
    END

```

```

      REAL*8 FUNCTION RFACT(VAL,POISON)
      REAL*8 VAL
      LOGICAL POISON

```

```

      RFACT=0.0D0
      IF (.NOT.POISON.AND.VAL.GT.1.5D0) THEN
        DO 100 N=2,INT(VAL+0.5D0)
          RFACT=RFACT+DLOG(FLOAT(N))
100 CONTINUE
      ENDIF
      RETURN
    END

```