MODELLING THE EFFECT OF NON-ACOUSTICAL VARIABLES ON INDIVIDUAL RESPONSE TO COMMUNITY NOISE

by

DESMOND McCAFFERTY, B.A.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Arts

McMaster University

September, 1978

MODELLING INDIVIDUAL RESPONSE TO COMMUNITY NOISE

MASTER OF ARTS (1978)
(Geography)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Modelling the Effect of Non-Acoustical Variables on Individual Response to Community Noise

AUTHOR: Desmond McCafferty, B.A. (University College, Dublin)

SUPERVISOR: Dr. F.L. Hall

NUMBER OF PAGES: ix, 151

ABSTRACT

This thesis is concerned with the effects of non-acoustical variables on individual response to noise.

Starting with some fundamental ideas about the nature of qualitative response, a specific mathematical model of individual response is derived. The model developed, a binary logit model, is suitable for the analysis of dichotomous (0, 1) response variables.

Two methods by which this model can be used to study the effect of respondent-specific, non-acoustical variables on response to noise are described and a two-stage analytical design which incorporates both approaches is developed.

The first part of the analysis combines characteristics of the respondent with noise level measurements in multivariate models of response. The second stage looks at differences in response across groups of individuals which are internally homogeneous with respect to certain key characteristics which influence response.

The general conclusions of the study are that (i) noise levels alone give a very poor explanation of the variance in individual response and (ii) the variables which influence response tend to be psychological in nature rather than socio-economic or demographic.

ACKNOWLEDGEMENTS

I wish to thank my supervisor, Dr. Fred Hall, for his advice and encouragement throughout the course of this study, and above all for his clear thinking on a number of confusing issues. I am also grateful to Dr. Michael Webber and Dr. Martin Taylor for their comments on an earlier draft of this thesis. Thanks are also due to Dr. Lee Liaw who advised on statistical matters.

Finally, I wish to pay tribute to Sharon Wright for her enthusiasm and willingness to type this thesis at extremely short notice.

TABLE OF CONTENTS

LIST OF TABLES

TABLE		Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In common with the other major pollutants, environmental noise is to a large extent the inevitable concomitant of social and economic advances. Given present levels of technology, the goal of noise abatement is incompatible with other major societal goals, the most obvious example being increased accessibility. Consequently noise - in particular transportation-related noise - is an important factor in transportation and land-use planning. It is, moreover, a factor of considerable complexity which not only pervades the physical and design aspects of planning, but which may also raise important issues of a social nature.

This thesis investigates the latter idea. The basic question which it addresses is To what extent does response to noise depend upon non-acoustical variables? Existing noise control ordinances have been drawn up on the implicit assumption that response is largely, if not completely, determined by the noise level alone; but there is an increasing body of evidence which suggests that characteristics of the individual are equally important in this respect. If personal variables *are* important, then predictions of response derived from noise levels alone may be inaccurate, and noise standards based upon these predictions ineffective.

1.1 Problem Statement

Existing studies shed little light on the effect of non-acoustical variables on response to noise. Few studies have even attempted to address the subject, tending instead to focus on the relationship between noise and response whilst ignoring the "intervening" (i.e. respondent-specific) variables. Typically, the focus of research in studies of response to both road traffic and aircraft noise has been the link between the average noise levels measured at residential sites and some measure of the average or median response, the latter usually measured according to subjective ratings of annoyance. In most cases the objective has been to develop a noise measure to correlate highly with annoyance scores (Griffiths and Langdon, 1968; Hazard, 1971; Rylander, Sorensen and Kajland, 1976), though occasionally the focus has been on the measures of response used (McKennell, 1970; Berglund *et al.*, 1974). However little progress has been made in these research areas in so far as there does not yet exist any widely accepted description of the noise-response relationship. Given that the majority of these studies have been conducted under largely uncontrolled conditions, this result is not surprising. Clearly, if socio-economic, demographic, personal or psychological characteristics of the respondents play any part whatsoever in affecting response to noise, failure to control for these factors from one study to the next will result in divergent findings. Langdon (1976) for example attributes the poor performance in reliability tests of the Traffic Noise Index developed by Griffiths and Langdon (1968), to differences between the samples used for developing and testing the index. If we are to achieve transferability with these models we need to be able to identify the effects of sample characteristics, and to broaden the specification of our predictive equations to include those characteristics significantly related to response.

A few studies have gone some way to providing an improved understanding of the role of respondent characteristics in determining response to noise (Langdon, 1976; Taylor and Hall, 1977a). However, the tests conducted by Taylor and Hall, though they cover a wide range of intervening variables, indicate only the level of significance of relationships between pairs of response variables and characteristics of the respondent. No attempt is made to develop a functional relationship between these components of the noise response problem. The Langdon study, on the other hand, does undertake such an analysis using a multiple regression model to specify functional relationships between dissatisfaction (the dependent variable), and noise level plus certain intervening variables. However only two intervening variables are examined, the respondent's subjectively rated satisfaction with his neighbourhood and his sensitivity to noise. In addition, the use of regression analysis with variables which are only ordinally measured is based upon the strong assumption of equal intervals between scale points: no attempt has been made, by Langdon or anyone else, to validate this assumption for the kind of data normally collected in social surveys on noise annoyance. The ordinality of the dependent variable in particular raises interpretational problems if the results are to be used for predictive purposes. It is the failure to identify an appropriate and statistically sound modelling technique for the analysis of individual-level data which is in part responsible for the lack of information on the influence of respondent characteristics on response to noise.

1.2 Research Objectives

The present study develops and applies an appropriate method for analyzing response to noise when the dependent variable has less than interval scale properties. The model used, a binary logit model, and the stimulus-response theory on which it rests were developed first in biological assay, though recent applications in travel demand modelling are probably more familiar to the geographer. With the exception of one recent rather cursory application (Starkie, 1975), this model has never been used to study response to noise. This is rather surprising in view of the fact that its theoretical/conceptual structure readily comprehends the noise response problem. Because the present study represents a relatively new area of application the derivation of the model from first principles is presented in some detail. The model is then used to analyze the effects of various socio-economic, demographic and personal characteristics on response to noise. Particular attention is paid to noise from main road (expressway) traffic. The thesis embodies two research objectives: first, to develop a model suitable for the analysis of individual response to noise and the effect of personal attributes on response; and second, to use this model to explore differences in reactions to noise across groups of individuals characterized by different levels of these intervening variables.

1.3 The Data Set

The data for this application comprise social survey information and sound level measurements collected at residential sites in the Hamilton-Toronto area over the summers of 1975 and 1976. The sample design and the data collected are described elsewhere (Taylor and Hall

1977b); accordingly no further description is included here. However a copy of the 1975 questionnaire is provided in Appendix 1.

1.4 Outline of the Thesis

The next chapter develops the model to be applied in subsequent analysis as one of a family of models appropriate for the analysis of qualitative response variables. Alternative forms and extensions of the model are treated, and reasons for the choice of the binary logit stated. Chapter 3 outlines the analytical framework for application of this model to noise response data, and describes the statistics used to guide the analysis. Following this are the two analysis chapters. Chapter 4 presents the results of some multivariate model estimations used to select intervening variables for more detailed analysis. In Chapter 5 the variables chosen are used to stratify the sample so that differences in responsiveness to noise across the different groups can be evaluated. Chapter 6 summarizes the results and concludes the thesis.

CHAPTER 2

DERIVATION OF THE MODEL

The model developed here for the study of individual response to noise is superficially similar to that used in travel demand forecasting, but its derivation is somewhat different. Because of two obvious contextual differences it is not possible to derive the noise model from similar theoretical constructs as those developed for example by Domencich and McFadden (1975). First, whereas it is usually possible in noise response studies to rank the responses according to the strengths of the reactions they index, this is not the case when dealing with travel-related choices. For example "extremely disturbed" is obviously a stronger response to noise than "slightly disturbed"; but it is clearly impossible to order typical mode split responses such as "choose auto" or "choose bus" using the same (or a similar) criterion. The result is that the multinomial model used to predict ranked responses to noise contains definitions of the various response probabilities which are structurally different from those used in multinomial travel choice models. But even where a response ranking is either impossible or undesirable the parameter estimation problem will still be different due to a second and more fundamental (though related) difference. This arises from the fact that while travel demand response categories or choice alternatives are individually characterized by particular values of the independent variables (time, cost or whatever), this is not so with categorized response to noise. Consequently it is not meaningful in dealing with noise to talk about the (dis)utility derived from making a response, so that utility maximization ideas are of no use in this context. It becomes necessary therefore to base the derivation of a noise response model on entirely different theoretical arguments and concepts. These are developed in the next section.

2.1 Conceptual and Theoretical Underpinnings

An individual's response to noise depends upon: (i) characteristics of the noise; and (ii) characteristics of the individual. Let Y denote the response variable and L some measure of the noise level which fully characterizes this stimulus. Denote by $\underline{\underline{Z}}$ a vector of observed characteristics of the individuals, which will in general include such variables as age, sex, income, occupation, etc. Finally $\underline{\underline{U}}$ is a vector of unobserved (or unobservable) variables which accounts for the residual variation in response unexplained by $\underline{\underline{Z}}$. The relationship between these components can be expressed as

$$Y = f(L, \underline{\underline{Z}}, \underline{\underline{U}}, e) \tag{2.1}$$

where e is a random error term due to measurement errors.

Assume initially that the individuals in our sample are relatively homogeneous with respect to the socio-economic and demographic variables represented by $\underline{\underline{Z}}$. Then response to noise is explained by the noise level L and the vector $\underline{\underline{U}}$, and so equation (2.1) simplifies to:

$$Y = f(L, \underline{\underline{U}}, e) \qquad (2.2)$$

If the response variable, Y, is measured on a continuous scale then analysis can proceed by using standard techniques such as regression analysis. Difficulties arise however if the response scale is quantal or qualitative in nature. Given that the scales used to measure noise response are usually ordinal and hence of the latter kind, the question which arises is whether response to noise is inherently "lumpy"; or whether there is a continuous (quantitative) scale of measurement underlying these qualitative scales. If the former is the case then clearly there must exist, for each individual, some threshold or critical noise level below which a particular response does not occur and at or above which it does (Finney, 1971). This critical value is usually referred to as a tolerance. Even if response is not quantal, but the measurement scale is, the idea of a threshold is still needed in order to relate the qualitative scale used to the underlying quantitativescale: that is, the use of mutually exclusive response categories implies thresholds corresponding to the boundaries of these categories. In short, the use of qualitative response variables involves, implicitly or explicitly, the assumption of tolerance levels. In the present study the assumption is made explicit and the construct is central to the theoretical development which follows - though the mathematical models which are derived do not depend upon it (Ashton, 1972).

Though the question of qualitative versus quantitative response is in some sense a moot point, two arguments support the hypothesis of quantal response. First, even when attempts are made to gain interval or ratio measurement of response there are doubts as to the success of such efforts. These arise partly because subjects tend to use only integers in response to questions designed to gain a continuous numerical rating of disturbance. Secondly, noise control regulations are founded upon the assumption of quantal response for otherwise there would be little point in setting such standards. Consequently it is assumed here that response to noise is in fact qualitative, so that the tolerance idea is used not as a by-product of a particular method of collecting data, but as an essential part of a theory which attempts to explain the phenomenon of quantal response.

Let us assume that only two mutually exclusive responses are possible: these can be denoted "disturbed" and "not disturbed". (This assumption will be relaxed in Section 2.4 to deal with the multiple response situation.) In this case the tolerance or threshold noise, denoted T_L, is that noise at which an individual becomes disturbed. This tolerance can be thought of as reflecting the combined effect of all the unobserved variables, including those which may be unique to a particular individual. Equation (2.2) can therefore be written as:

$$Y = f(L, T_L, e) \tag{2.3}$$

By further assuming that the noise level and response are measured without error (an assumption retained throughout), e disappears, yielding

$$Y = f(L, T_L) \tag{2.4}$$

An individual will be disturbed if the actual noise exceeds his tolerance for noise, i.e., if ($T_L < L$). But T_L is an unobservable, so that we do not have any information about the occurrence of this event. However by making assumptions about the distribution of T_L in the population we can at least answer questions about the probability of the event, for each individual, given of course the actual noise level to which each is exposed.

2.2 Linear, Probit and Logit Models of Response

Following Domencich and McFadden (1975) we can identify at least three different assumptions about the distribution function of T_L which are of interest either because of their intuitive appeal or because they yield models which are easy to work with.

Assumption 1: Suppose that T_L has a uniform distribution on the interval (a, b). Then:

$$P = P(T_L < L) = F(L) = \begin{cases} 0 & L \leq a \\ \int_a^L \frac{1}{b-a}\, dt & a < L < b \\ 1 & L \geq b \end{cases} \qquad (2.5)$$

where P is the probability that a person exposed at noise level L will be disturbed. Performing the integration in (2.5) we obtain for $a < L < b$:

$$P = (L-a)/(b-a)$$

$$= -\frac{a}{b-a} + \frac{1}{b-a} L \qquad (2.6)$$

The result therefore is a truncated linear probability model (Figure 2.1) with parameters $\alpha = -a/(b-a)$ and $\beta = 1/(b-a)$.

Assumption 2: Alternatively we might assume that T_L is normally distributed $N(\mu, \sigma)$ throughout the population. Hence

$$P = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{L} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right)dt$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{X} \exp\left(-\frac{s^2}{2}\right)ds \qquad (2.7)$$

Here $X = \frac{1}{\sigma}(L-\mu)$ is the normit of P, being the value on the abscissa which yields a probability of P in a standard normal distribution. Simply in order to avoid negative values of X it is usual to write equation (2.7) as:

$$P = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{X-5} \exp\left(-\frac{s^2}{2}\right)ds \qquad (2.8)$$

where $X = 5 + \frac{1}{\sigma}(L-\mu) = (-\frac{\mu}{\sigma} +5) + \frac{1}{\sigma} L$ is known as the probit of P (Finney, 1971). Under the probit transformation therefore P is the area under a normal N(5, 1) curve on the interval $(-\infty, (-\frac{\mu}{\sigma} +5) + \frac{1}{\sigma} L)$ (Figure 2.2). The parameters of interest in this model are $\alpha = -\frac{\mu}{\sigma} +5$, $\beta = \frac{1}{\sigma}$.

Assumption 3: Finally we might assume that T_L has a logistic distribution. This gives:

$$P = \frac{1}{1 + \exp(-\alpha-\beta L)} \qquad (2.9)$$

Figure 2.1

Truncated Linear Probability Model

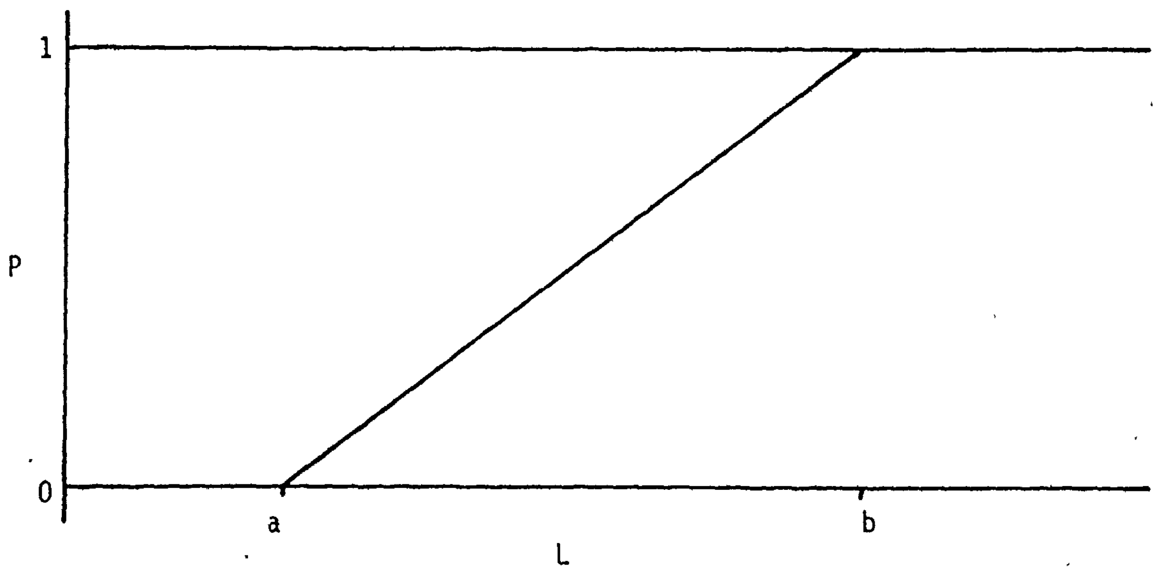
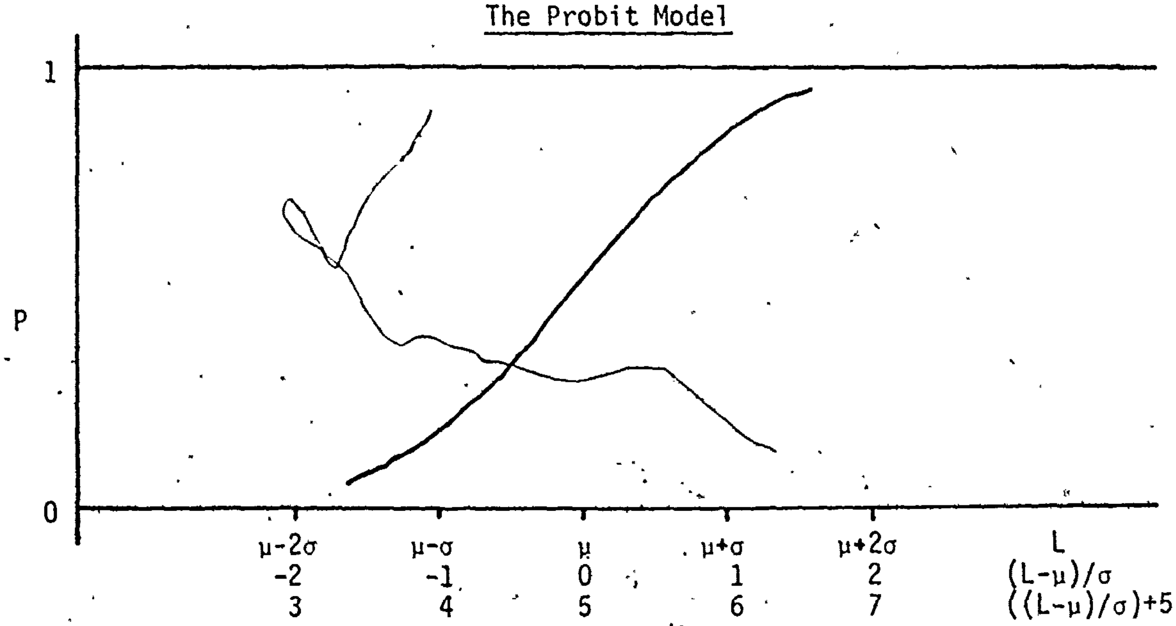


Figure 2.2

The Probit Model

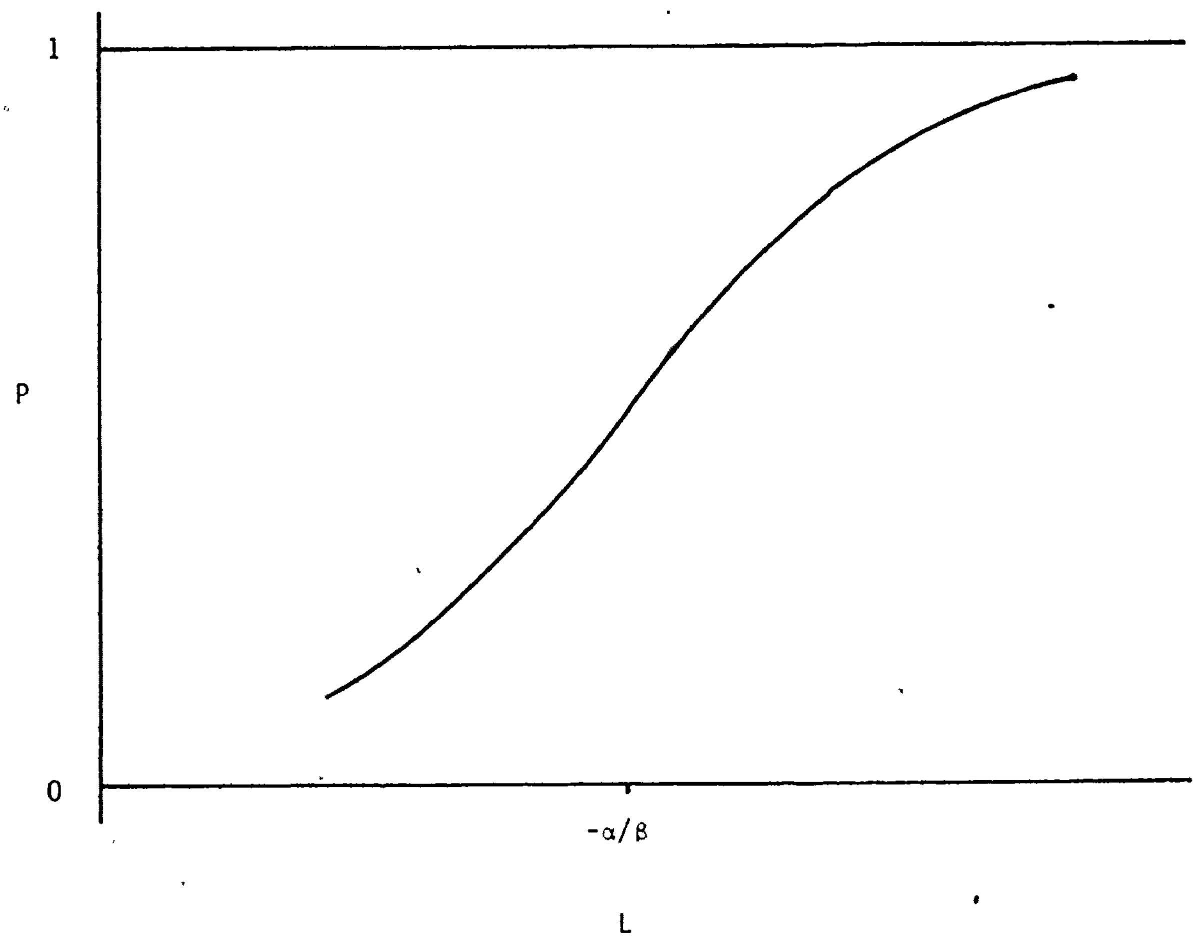


which defines a logit model with parameters α and β (Figure 2.3). The logistic is a symmetrical distribution widely used in place of the normal when a simpler distribution is needed. In fact the percentage points of the standard logistic distribution closely approximate those of the standard normal though there is some deviation in the probabilities for extreme values of L (Domencich and McFadden, 1975).

As regards the selection of an appropriate model from these three, the linear probability model can probably be dismissed first on the grounds that the assumption of a uniform tolerance distribution is implausible in most stimulus-response situations. Moreover, as Domencich and McFadden (1975), have shown, the presence of observations for which L is less than a or greater than b, will result in the parameter estimates being biased, perhaps substantially, *below* their true values. The distribution assumptions underlying both the logit and probit models are more tenable, but unfortunately because of the nature of T_L they cannot be subjected to empirical verification. There are therefore no strong grounds on which to discriminate among these two, so that choice between them is usually guided by practical considerations more than anything else. In this respect the logit model has usually been favoured to date by reason of its more tractable form. The problem with the probit model is that the probability P defined in equation (2.8) cannot be estimated by simpler functions so that probit analyses by computer have to draw on tables of the percentage points of the normal distribution. In the present case the logit model is chosen as the analytical tool because of the practical advantage offered by its relatively simply functional form.

Figure 2.3

The Logit Model



2.3 The Multivariate Logit Model

The assumption of a population which is homogeneous in everything except the noise level affecting individuals is neither realistic, nor, in the present context very helpful, since we are particularly interested in the effect of variation in $\underline{\underline{Z}}$ across the sample. Relaxing this assumption we obtain, again assuming absence of measurement error:

$$Y = f(L, \underline{\underline{Z}}, \underline{\underline{U}}) \qquad (2.10)$$

The development of a model of response for a population which is heterogeneous with respect to certain variables which influence response involves retention of the tolerance concept. However whereas previously $\underline{\underline{U}}$ was the only source of variation among individuals, and as a result individual tolerances depend upon $\underline{\underline{U}}$ only, the more general case with which we are now dealing requires that:

$$T_L = f(\underline{\underline{Z}}, \underline{\underline{U}}) \qquad (2.11)$$

Clearly, the distribution of T_L in the population will be affected by the distributions of the variables $Z_1, \ldots, Z_K$. The effect of these variables on response may be examined in either of two ways. First we may stratify the sample so as to obtain groups of individuals which meet the assumption of a constant $\underline{\underline{Z}}$ vector. Within each of these groups variation in tolerance levels is due solely to variation in the unobserved variables $\underline{\underline{U}}$ and so analysis can proceed as described above by assuming a distribution for T_L within each group. Provided the assumptions (models) are the same for the various groups we can then test for differences in the parameters of the response equation across these groups.

Alternatively we can develop a multivariate response model which explicitly incorporates the vector of observed variables, $\underline{\underline{Z}}$. For illustrative purposes, suppose that $\underline{\underline{Z}}$ contains one variable only, the age of the respondent, which we shall assume to be negatively related to disturbance as a result of presbycusis-induced increase in the individual's tolerance for noise. Since T_L varies directly with age, the probability of disturbance must now be considered in relation to a *conditional* distribution of tolerances - the distribution of tolerance levels given age. A series of age-conditional tolerance distributionsis illustrated in Figure 2.4 in which the distributions for successive ages are shifted upwards along the tolerance axis as a result of a constant or generic increase in the individual (and hence the average) tolerance for noise.

Formally, let A denote the random variable age, and f(A) some function thereof, then assume:

$$T_{L|A} = T_{L|f(A)=0} + \theta \cdot f(A) \tag{2.12}$$

where: $T_{L|A}$ = conditional tolerance of individual given age

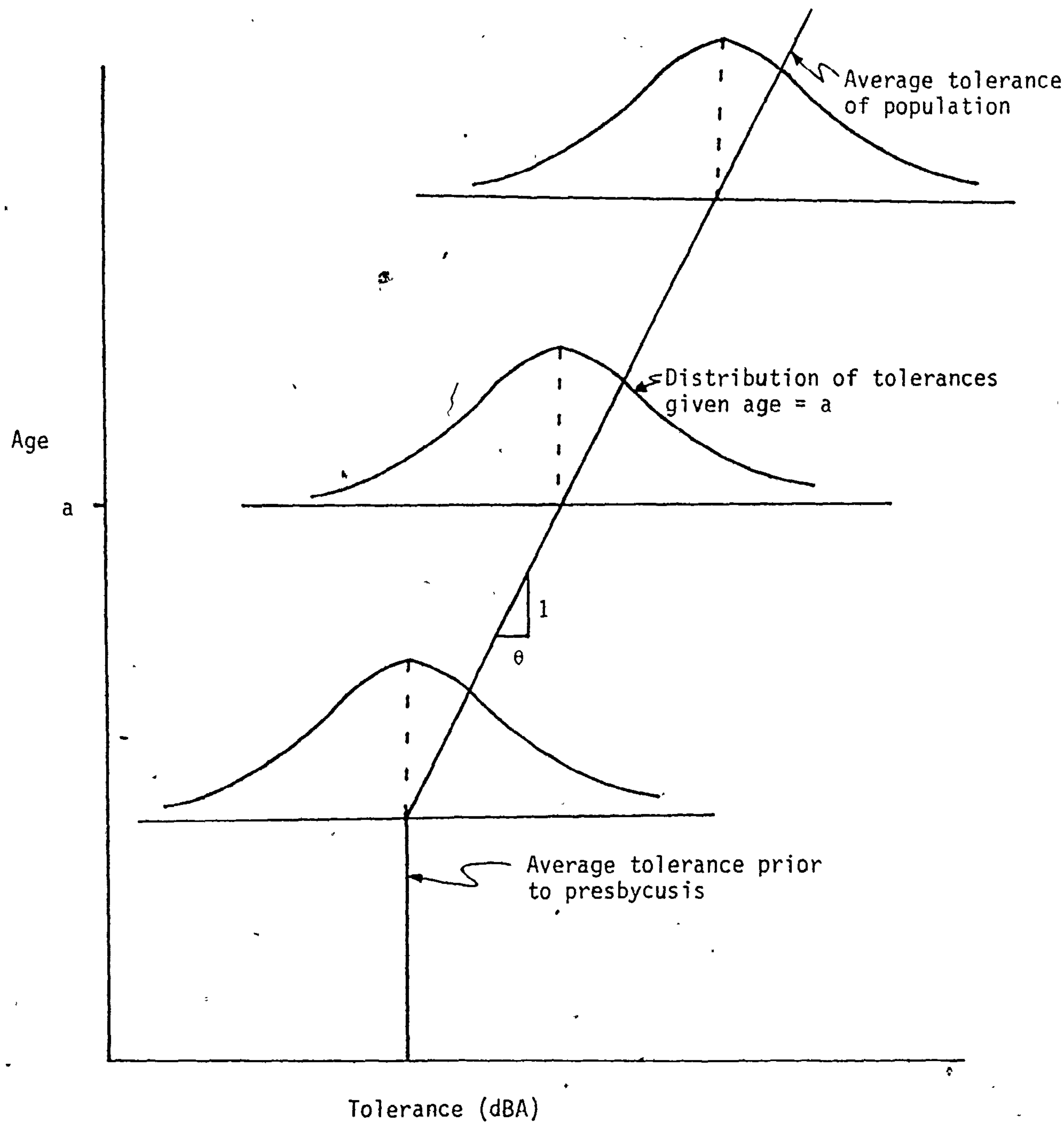
$T_{L|f(A)=0}$ = the initial (prior to presbycusis) tolerance of the individual

θ = a parameter to be determined empirically.

Equation (2.12) states that tolerance is a linear-in-parameters function of age over the region bounded by f(A)=0. The function f(A) could simply be the identity transformation, but a non-linear form is more likely, if for no other reason than that the effect of this variable on tolerance may be zero over a particular range of ages. In terms of the present example,

Figure 2.4

Conditional Distribution of Tolerance Given Age



presbycusis may be absent below a particular age. However, it may not always be the case that we can so conveniently identify a range of a variable over which its effect on tolerances is zero, due to the lack of any well-developed explanation of causative links such as that provided in the case of age by the concept of presbycusis. Variables such as Income and Education are obvious examples. In these cases we can instead identify some value(s) of the variable, Z, for which f(Z) is an extremum and which can therefore be regarded as a base condition. $T_{L|f(Z)=0}$ can therefore be interpreted as the tolerance of a "control" group of individuals.

An individual will be disturbed if

$$T_{L|A} < L$$

that is, if

$$T_{L|f(A)=0} + \theta \cdot f(A) < L$$

or

$$T_{L|f(A)=0} < L - \theta \cdot f(A) \tag{2.13}$$

To establish the probability of disturbance we need to make an assumption about the distribution of $T_{L|f(A)=0}$. Assume this to be logistic: then

$$P(T_{L|A} < L) = \frac{1}{1 + \exp(-\alpha-\beta_0(L-\theta\cdot f(A)))}$$

$$= \frac{1}{1 + \exp(-\alpha-\beta_0 L-\beta_1 f(A))} \tag{2.14}$$

where: $\beta_1 = -\beta_0 \cdot \theta$.

Generalizing equation (2.14) to include the entire vector of attributes $\underline{\underline{Z}}$ yields

$$P = P(T_{L|\underline{\underline{Z}}} < L) = \frac{1}{1 + \exp(-\alpha-\beta_0 L-\beta_1 X_1-\ldots-\beta_K X_K)} \quad (2.15)$$

where X_K is some empirical function of Z_K. To illustrate: if a particular variable X_k is negatively related to tolerance, then θ_k will be negative and β_k positive, so that, as we would expect, the probability of disturbance will increase with X_k.

Equation (2.15) represents the form of the multivariate logit model commonly used. What has not previously been recognized however is the assumption, implicit in equation (2.12), that the conditional distributions of tolerance at each value of $\underline{\underline{Z}}$ are identical but for the locational parameter (c.f. Figure 2.4): that is, the coefficient of noise in equation (2.15), β_0, is constrained to remain constant regardless of the value of $\underline{\underline{Z}}$. In contrast, the stratification method of analyzing the effect of $\underline{\underline{Z}}$ on response allows variation across groups (ages) in both the location *and* scaling parameters of the tolerance distribution, and is therefore based upon a less restrictive assumption about the nature of the relationship between $\underline{\underline{Z}}$ and T (equation 2.11).

2.4 Multinomial Extensions

So far only the case of a dichotomous (0, 1) dependent variable has been treated. We next consider the extension of the binary logit model developed above, to the situation in which more than two responses are possible. Two rather different versions of the logit model have been proposed as suitable for the analysis of multi-category response variables. The difference between these is that one assumes the dependent variable can be related to at least an ordinal scale of measurement; the other does not (Cox, 1970). Both models are developed here for the simplest case in which there is a single explanatory variable but again the multivariate generalization should be obvious.

The first model considered applies when there exists some kind of ordering among responses. For example respondents may be asked to rate the level of noise in their neighbourhood as being either 'not at all disturbing', 'disturbing' or 'highly disturbing'. These response categories might be ranked 0, 1 and 2 respectively, using strength of response as the ordering rule. In general there will be J+1 response categories, ranked 0 through J. We can now extend the arguments presented above for binary response variables by positing that for each individual there exists a set of threshold noise levels $\{T_j: \ j = 1, \ldots, J\}$, where T_j is the level of noise above which the individual makes response j and below which he makes response j-1 (the subscript L is omitted for clarity). The probability that a given individual will make response j, for $j = 0, \ldots, J$ is given as:

$$P_J = P(T_J < L)$$

$$\begin{aligned} P_{J-1} &= P(T_{J-1} < L \leq T_J) \\ &= P(T_{J-1} < L) - P(T_J < L) \\ &\quad \vdots \qquad\qquad \vdots \end{aligned} \tag{2.16}$$

$$P_1 = P(T_1 < L \le T_2)$$
$$= P(T_1 < L) - P(T_2 < L)$$

$$P_0 = P(L \le T_1) = 1 - P(T_1 < L)$$

If we assume that T_j is logistically distributed with parameters α_j, β, for $j = 1, \ldots, J$, then these probabilities can be given explicit functional forms:

$$P_J = (1 + e^{-\alpha_J - \beta L})^{-1}$$

$$P_{J-1} = (1 + e^{-\alpha_{J-1} - \beta L})^{-1} - (1 + e^{-\alpha_J - \beta L})^{-1}$$
$$= (1 + e^{-\alpha_{J-1} - \beta L})^{-1} - P_J$$

$$P_{J-2} = (1 + e^{-\alpha_{J-2} - \beta L})^{-1} - (1 + e^{\alpha_{J-1} - \beta L})^{-1}$$
$$= (1 + e^{-\alpha_{J-2} - \beta L})^{-1} - (P_{J-1} + P_J)$$

$$\vdots \tag{2.17}$$

$$P_1 = (1 + e^{-\alpha_1 - \beta L})^{-1} - (P_2 + \ldots + P_J)$$

$$P_0 = 1 - (P_1 + \ldots + P_J)$$

More generally,

$$P_j = \begin{cases} (1 + e^{-\alpha_j - \beta L})^{-1} & j = J \\ (1 + e^{-\alpha_j - \beta L})^{-1} - \sum_{i=j+1}^{J} P_i & 0 < j < J \\ 1 - \sum_{i=1}^{J} P_i & j = 0 \end{cases} \quad (2.18)$$

The logic of the problem requires that the parameter β be held constant while

$$\alpha_j > \alpha_{j+1} \qquad j = 1, \ldots, J-1 \qquad (2.19)$$

(Gurland, Lee and Dahm, 1960). To see this consider the example cited above in which there are 3 possible responses, denoted 0, 1 and 2. The cumulative distribution curves of T_1 and T_2 intersect when

$$(1 + e^{-\alpha_1 - \beta_1 L})^{-1} = (1 + e^{-\alpha_2 - \beta_2 L})^{-1}$$

That is, when

$$\alpha_1 + \beta_1 L = \alpha_2 + \beta_2 L$$

or,

$$L = \frac{\alpha_1 - \alpha_2}{\beta_2 - \beta_1} \qquad (2.20)$$

If $\beta_1 = \beta_2$ equation (2.20) is undefined, i.e. no solution exists, which means the curves will not intersect. However, if $\beta_2 \neq \beta_1$ the point of intersection exists, and implies the existence of a value of L for which

$$(1 + e^{-\alpha_1 - \beta_1 L})^{-1} < (1 + e^{-\alpha_2 - \beta_2 L})^{-1}$$

which, using (2.17) implies that

$$(P_1 + P_2) < P_2$$

for a contradiction. In turn it can be seen that for β constant, (2.19) is a necessary and sufficient condition for the consistency of the model.

In general then, where there are K independent variables, this model uses information on J+1 possible responses to generate the J+K parameters necessary to fit a series of J parallel response curves to the data.

The second model applies whenever there is no obvious ranking of the alternatives, or where we wish to use a more general model that takes no account of the ordering. As an example, Wrigley (1976) cites the situation where a sample of respondents indicate whether they are 'in favour', 'against' or 'undecided' about a proposed road improvement scheme. These responses can be *arbitrarily* scored 2, 1 and 0 respectively. Because of the fact that the dependent variable in this case has only nominal properties, the concept of a tolerance level governing transition (progression) from one response to the next is useless. The development of the model for this situation has been based instead on contingency table analysis (Cox, 1970; Mantel, 1966), though Wrigley's derivation follows from what he appears to consider an intuitive generalization of the binary logit model. Both approaches at some point designate paired parameters α_j and β_j to correspond to each of the possible responses. Then the log odds of being in category j as opposed to category j+1 is given by

$$\log_e\left(\frac{P_j}{P_{j+1}}\right) = (\alpha_j - \alpha_{j+1}) + (\beta_j - \beta_{j+1})L \qquad (2.21)$$

which is equivalent to

$$P_j = \frac{e^{\alpha_j+\beta_j L}}{\sum_{j=0}^{J} e^{\alpha_j+\beta_j L}} \qquad (2.22)$$

In order to obtain unique parameters (only the differences in (2.21) matter), it is usual to impose the constraint that

$$\alpha_0 = \beta_0 = 0$$

In this case then there are J(K+1) parameters to be estimated, where again K is the number of independent variables.

In most noise response studies conducted to date the dependent variable has been assumed to have ordinal properties. This assumption has also been made in previous work using the dependent variables to be analyzed here (Taylor and Hall, 1977a), and is probably quite reasonable. This makes the ranked responses model the appropriate analytical tool to use both in the present context and in noise response studies in general (though Mantel 1966 offers a dissenting opinion). However, there would appear to be some question as to the value of a multinomial model since the additional information which it yields hardly justifies the extra computation involved in its estimation. A multinomial analysis would be a reasonable procedure if it could be shown to use more information in order to yield better estimates of a fixed number of parameters. These models use more information to generate *more* parameters rather than *better* parameters. Indeed the number of parameters in the unranked responses model is so large as to raise interpretational difficulties. For these reasons the simpler binary model is chosen for the present study.

2.5 Summary

This chapter has developed a formal model of response to noise. The theoretical foundation of the model is predicated upon the idea of a tolerance (threshold or critical) noise level which is held to account for the qualitative or quantal nature of individual response. By making assumptions about the distribution of tolerance levels in the population we can derive explicit mathematical forms for the response model. The logit model results from an assumption of logistically distributed tolerances. This particular postulate is neither more nor less realistic than others such as that of a normal tolerance distribution: it is preferred here simply because it gives rise to a model which is easy to calibrate. The tolerance concept can be used to deal with any number of response classes, but the multicategory model is not so much an extension or improvement of the simple two category case as structurally a different model. The more complex model provides more information but of a less useful kind and so it was decided to focus only on dichotomous response in the present study. This decision completes the specifications necessary to fully determine the mathematical form of the model - binary logit - to be used in subsequent analysis.

CHAPTER 3

SPECIFICATION OF VARIABLES AND ANALYTICAL DESIGN

The purpose of this chapter is to describe in detail the framework within which the model of response developed above is to be applied. Section 3.1 relates the abstract components of the model to specific empirically-measured variables. Methods by which these variables may be examined are discussed in the following section, and the analytical algorithm presented. The penultimate section, 3.3, describes the statistics used to guide the analysis and finally section 3.4 summarizes the chapter.

3.1 Variables for Analysis

In order to operationalize the model of response described in Chapter 2, we need to identify the variables which correspond to the stimulus and response components of the model. With regard first to the selection of a noise measure, the present trend in acoustics seems to be towards the use of relatively simple noise metrics and away from the more sophisticated measurement procedures (Schultz, 1972). In support of this movement Langdon (1976) found that elaborate equations such as that for the Traffic Noise Index (TNI) performed no better than the simpler measures such as L_{10}, L_{50}, L_{eq} etc., in explaining dissatisfaction scores. The measure used in the present analysis, the day-night equivalent sound level,

L_{dn}, is an L_{eq}-based index in which the nightime noise energy is multiplied by 10 and combined with the daytime level to give a weighted average for the 24 hour period, that is

$$L_{dn} = 10 \log[\frac{1}{24}(15L_d + 9(10L_n))]\text{dBA} \qquad (3.1)$$

where: L_d = average daytime (07:00-22:00) sound energy

L_n = average nightime sound energy.

Because it places more emphasis on nightime sound levels this measure has been identified as an appropriate basis for noise standards for residential areas (USEPA, 1974). It is used in the present context for the same reason and also to facilitate comparison with the results of other studies which have adopted L_{dn}.

Response measures are of two basic kinds - cognitive and behavioural. *Cognitive* measures evaluate an individual's reaction in terms of the formation of, or change in, attitudes and/or other psychological variables. Disturbance, dissatisfaction and annoyance are the names applied to the kinds of cognitive response measures commonly used in noise studies. Measures of *behavioural* response, on the other hand, assess the impact of noise (or any other stimulus) according to the physical actions or reactions of the individual. Complaint action is one of the most common measures of behavioural response used, largely because of its political importance.

In the present study both types of response measure were available from the data set. Cognitive response was measured according to the individual's rating of the disturbance caused him by noise. Other variables recorded what complaint action, if any, had been taken, as well as the incidence of both immediate reactions (e.g., closing windows) and long term reactions (e.g., changing residence) to the noise. For the purpose of this analysis however it was decided for a number of reasons to focus on cognitive response only. In the first place, several studies have shown that action taken, and in particular complaint activity, is an unreliable and unstable measure of the annoyance caused by noise (McKennell, 1970; Taylor and Hall, 1977a). Because the level of, for example, complaint activity depends upon certain social and political variables such as membership of organizations and political activity (McKennell, 1970) it has been suggested (Borsky, 1970) that a change in socio-political conditions could lead to a sharp increase in the number of complaints filed, *even if the noise level were to remain the same*. Indeed Bauer (1970) believes that this may already be happening. Secondly, cognitive response measures are chosen for the same reason that L_{dn} was selected as the noise metric, namely to facilitate comparisons with the results of other studies.

The actual measures of cognitive or psychological response which are used, consist of disturbance ratings for both neighbourhood noise in general and main road traffic noise in particular (Table 3.1). The reason for looking at response to overall noise is that many of the sites in the 1975 data set experienced high noise levels from sources other than road traffic, e.g. railway and industrial noise. Early approaches to the measurement of annoyance or disturbance used indirect techniques, constructing disturbance scales from the response to a host of questions dealing with activity interruption and other adverse noise impacts. However recent studies have tended to use the simpler and more direct method of obtaining

Table 3.1

Variables Used in Analysis

Variable	Name
Noise Metric	
Day-night equivalent sound level	L_{dn}
Response Variables	
Attitude towards overall neighbourhood noise - any disturbance	NHDATT (1-5/6-9)
Attitude towards overall neighbourhood noise - severe disturbance	NHDATT (1-7/8-9)
Attitude towards main road traffic noise - any disturbance	MNATT (0-5/6-9)
Attitude towards main road traffic noise - severe disturbance	MNATT (0-7/8-9)
Intervening Variables	
Age	AGE
Sex	SEX
Level of Education	LEVED
Income	INCOME
Tenure	TENURE
Length of Residence	LENRES
Hours spent at home during the week	WKHME
Hours at home during the weekend	WENDHM

Table 3.1 (cont'd)

Variable	Name
Sensitivity	SENSE
Rating of Workplace Noise	RATWORK
Sleep Interruption	SLPINT
Relaxation Interruption	RLXINT
Conversation Interruption	CONINT
Work Interruption	WKINT
T.V. Interrupted	TVINT
Phoning Interrupted	PHINT
Eating Interrupted	EATINT
Shielding (presence or absence of a wall)	WALL

subjective ratings of disturbance. Langdon (1976) presents evidence to suggest that elaborate scaling procedures offer no advantage from the point of view of correlation with measured noise levels.

The scales used to rate disturbance in the present study are 9-point bipolar scales, with the extreme points labelled as "extremely agreeable" and "extremely disturbing". Intervening points were labelled to correspond to degrees of disturbance intermediate between these extremes (see question 6 of questionnaire in Appendix 1).

Since the model selected in Chapter 2 requires the dependent variable to be in binary form, these scales have to be dichotomized. This requires the selection of a cut-off point which divides the scale into two general response categories. In the present case two dichotomizations of the scale were of interest - one separating those individuals who reported *any* disturbance from those not at all disturbed; and the other differentiating between the "highly disturbed" respondents and those with lower disturbance levels. The former requires a cut-off directly above the neutral point on the scale (i.e., neutral is included with the "no disturbance" category), and is a fairly obvious dichotomization to try. For the second test, "highly disturbed" individuals were defined as those rating the noise as either "considerably disturbing" or "extremely disturning" (Appendix 1). Schultz (1977) has argued strongly in favour of the latter classification system, claiming that it is a more sensitive measure of the adverse impact of noise, and hence more suitable for regulatory purposes.

In summary then, we have, in all, four distinct dependent variables which together cover response to both main road traffic and overall neighbourhood noise levels, where response is defined according to two different criteria.

The battery of intervening variables (Table 3.1) used in conjunction with noise level measurements (L_{dn}) to explain response includes fairly standard socio-economic variables such as income and education, as well as such demographic variables as age and sex. Other variables measure the length of residence and tenure characteristics of the household while two measures of the time spent at home by the respondent are also included for obvious reasons.

In addition to these variables are measures of the activity interruption caused by noise and two psychological characteristics of the individual - his reported sensitivity to noise and rating of the noise level at his place of work. Both of these variables are measured on labelled 5-point bipolar scales. Sensitivity is included on the strength of Langdon's (1976) finding that this variable accounted for a large part of the variation in individual response. The rating of workplace noise is the kind of attitudinal variable that McKennell (1970) suggests is likely to be closely related to subjective ratings of annoyance.

Activity interruption is measured for sleeping and conversation as well as a number of other daily activities. It is possible to view interruption measures as valid indices of the adverse impact of noise and so to treat them as dependent variables to be explained by noise exposure plus other independent variables. However because it was decided, as described above, to use a more direct measure of annoyance these measures are more appropriately viewed as explanatory variables since activity interruption is usually considered to occur prior to psychological response (Borsky, 1970; Taylor, Gertler and Hall, 1978). It is acknowledged, however, that there may be some question as to whether a set of explanatory variables which includes activity interruption measures along with variables such as sensitivity satisfies mutual independence requirements. We will return to this point later.

Finally the presence or absence of a wall as a shield between the respondent and the noise source is recorded. Earlier work on the present data set (Hall, Birnie and Taylor, 1978) suggests that for a given level of noise, respondents at sites with this type of shielding are more likely to be disturbed (by overall noise levels at any rate) than if no shielding at all, or some other form of shielding existed: that is, there appears to be some negative psychological effect associated with this type of barrier working in opposition to its acoustical effect.

3.2 Analytical Procedures

Two methods of looking at the effect of intervening variables on response have been described already (Chapter 2, Section 2.3). One of these involves the use of a multivariate logit model: the other is based upon a stratification of the sample for the purpose of comparing models calibrated for the resulting groups. The second (stratification) approach is superior in several respects. Most importantly, it is based upon less restrictive assumptions about the relationship between various components of the model of response developed above (c.f., Section 2.3). As a result it distinguishes clearly between stimulus and intervening variables whereas the multivariate model described by equation (2.15) confuses these components. The stratification approach is therefore easier to relate to the conceptual model of response developed by Borsky (1970), and McKennell (1970), and more directly relevant to the aims of this thesis. The major disadvantage of this procedure is that it is more time consuming than direct multivariate analysis if the number of intervening variables is large. Under such circumstances there may also prove to be too few observations in certain cells of the attribute matrix, particularly if the sample was not based upon a comprehensive stratification design. The result is that it may not be feasible to hold all variables constant at once.

In order to gain the advantages of both approaches the analytical design for this thesis is based upon a two-stage analysis, as described next.

3.2.1 Outline of the Multivariate Analysis

In the first stage the intervening variables listed in Table 3.1 are used in conjunction with the noise level to build up the best multivariate model for each response variable. This analysis serves a twofold purpose: first it specifies the functional relationships between the intervening variables and various measures of response; secondly it serves to narrow the focus of subsequent analysis by indicating the intervening variables which are significantly related to response.

The first problem encountered in this part of the analysis lies in the fact that most of the intervening variables listed in Table 3.1 are either ordinal or nominal in measurement and cannot be directly incorporated into equation (2.15) without some prior transformation. For a particular variable this consists of first identifying, *on the given measurement scale*, ranges of the variable or classifications of its possible values, within which the explanatory effect of that variable is hypothesized to be constant (Tulloch and McMillan, 1973). For variable Z_k, let there be r_k such classifications or groups, labelled $Z_{k1}, \ldots, Z_{kj}, \ldots, Z_{kr_k}$, where Z_{kj} is the jth classification of the variable. On the basis of these classifications the original variable Z_k is now transformed into a series of dummy variables $X_{k1}, \ldots, X_{kr_k}$ with:

$$X_{kj}^{(i)} = \begin{cases} 1 \text{ if individual } i \in Z_{kj} \\ 0 \text{ otherwise} \end{cases} \quad (3.2)$$

In order to avoid problems of multicollinearity we can designate a particular classification as a "base" group: the remaining r_k-1 dummy variables are then used to represent the variable in equation (2.15). Thus, if there are M non-interval scaled variables to be considered we have

$$P = [1 + \exp(-\alpha - \beta_0 L - (\sum_{k=1}^{M} \sum_{j=1}^{r_k - 1} \beta_{kj} X_{kj} + \sum_{k=M+1}^{K} \beta_k Z_k))]^{-1} \quad (3.3)$$

Here β_{kj} can be interpreted as a measure of the degree to which individuals in class Z_{kj} are more likely to respond than individuals in the base class who possess otherwise identical characteristics. Similarly the term $(\beta_{kj} - \beta_{kl})$ provides a measure of the difference in responsiveness of individuals in any two classes, Z_{kj} and Z_{kl} of the variable.

The obvious problem with this procedure is that the effect of many variables on response will be continuous, so that even within a particular class of the attribute, the probability of response will differ according to the "true" or underlying value of the variable. The justification for this method of handling certain variables derives largely from the fact that it is in some sense a *sine qua non* given that much of the data are only available from the questionnaire in ordinal or nominal form. This is not to suggest that the questionnaire is in any way lacking. Many of the variables are inherently nominal (e.g., sex or tenure) so that the above criticism does not apply. Others, for example income and the psychological variables such as sensitivity, would likely be subject to large reporting or measurement errors if an attempt were made to obtain interval scaled data, with the result that the data would only be reliable anyway when grouped. For this reason the handling of these variables in the manner described above seems reasonable.

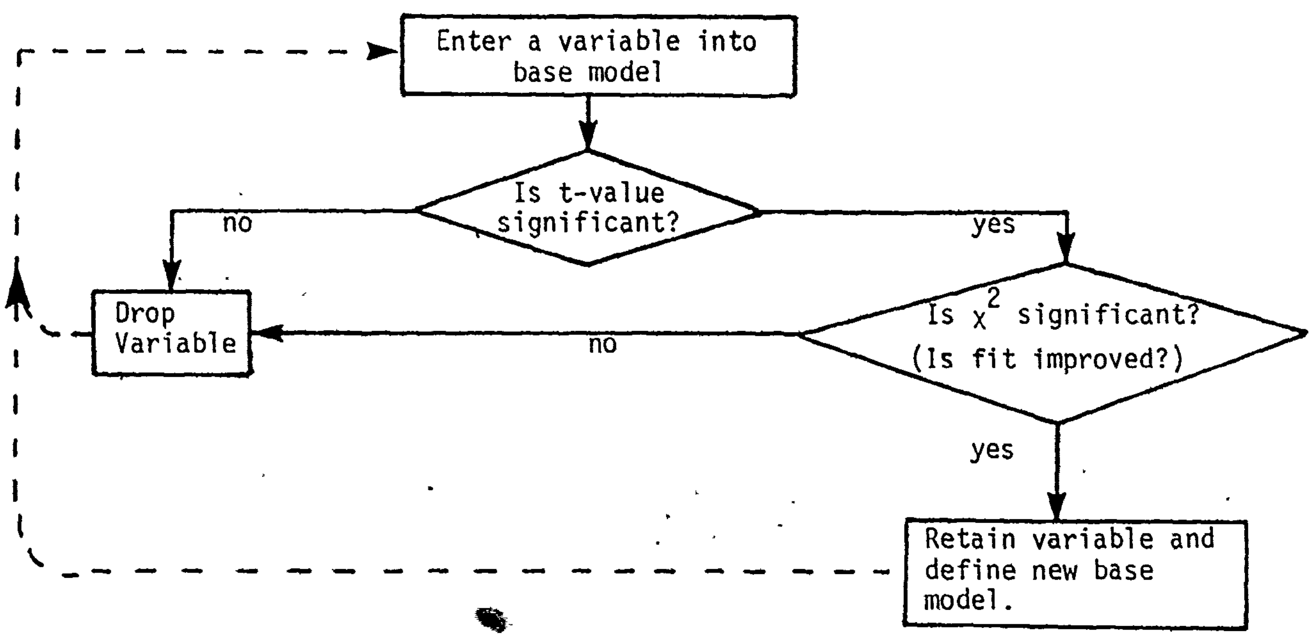
As is evident from equation (3.3) the expression for P can involve a very large number of terms when the number of variables, K, is large and contains variables for which r_k is also large. In fact because of this and the large sample size (1786 observations in all) storage problems were encountered in the computer analyses. In order to obviate these problems r_k was set equal to 2 for all the multicategory (ordinal) variables: that is, it was assumed that the differences in response *due* to a particular variable could be fully explained by assigning individuals to one of two groups according to their scores on that variable. The danger in this simplification is that it may obscure bimodal relationships. However preliminary analyses indicated that in all but a few cases relationships between these independent variables and all four of the dependent variables were monotonic. Moreover, in those cases where a more complex pattern of association existed, the relationships were not in any case significant.

This simplified the analysis somewhat but necessitated a further decision as to the definition of the two groups on each variable. As an example, suppose that income is hypothesized to be monotonically related to response such that higher-income individuals are more likely to respond than those with lower incomes; the question which arises, if only two income groups are to be compared, is where to place the cut-off between these groups. Empirical analysis was used to answer this question. Taking each of the four dependent variables in turn, the model was estimated for each intervening variable plus a constant. For a particular intervening variable every possible cut-off point on the measurement scale was tested, and that which gave the highest t-statistic for the variable in question as well as the best overall fit to the data was selected.

The first response variable analyzed was attitude to overall neighbourhood noise as described by the two response classes "not at all disturbing" and "disturbing". Analysis proceeded by designating the model containing only noise level (L_{dn}) and a constant a "base" model. Each intervening variable was then entered into this base model, in arbitrary order, and retained only if *both* its coefficient *and* contribution to overall fit were significantly different from zero (Figure 3.1). This algorithm was also used to analyze the neutral cut-off dependent variable for attitude to main road traffic noise. For the remaining two response variables (those defined by the more extreme cut-off point) the procedure was slightly different. In order to save time, both of these variables were first cross-

Figure 3.1

Algorithm for Analysis, Part 1



tabulated with each of the intervening variables, and χ^2 and ϕ measures of association computed for the resulting 2×2 contingency tables (Nie *et al.*, 1975). The base model was then formed from those independent variables which were significantly related to the response variable in question, as determined from the χ^2 statistics, *and* which had entered into the optimum model for the corresponding dependent variable defined according to a neutral cut-off point. The algorithm outlined in Figure 3.1 was followed for each of the remaining intervening variables which gave significant χ^2 statistics. The results of this stage of the analysis are reported in Chapter 4.

3.2.2 Stratification Design

The second stage of the analysis consists of stratifying the sample according to the intervening variables identified in stage 1 as significantly related to response. For consistency with the first part of the analysis it is necessary to look at these variables in their reduced (dichotomized) form using the same cut-off points as before. For each response variable the sample is stratified according to the q intervening variables significantly related to it. The stratification matrix therefore contains 2^q cells. The effect of any one variable, all others being held constant, may then be examined by comparing any of 2^{q-1} pairs of cells defined along the appropriate axis of the q-dimensional matrix. In the present study the pair which was chosen from among these was simply that which had the largest number of observations in each of the two cells.

To be able to discuss differences in response across various groups of the population we need some statistic(s) on which to base comparisons.

Since the variables which enter the model calibrated for each group are the same in all cases (i.e. L_{dn} plus a constant) the focus of attention should be the differences in the coefficients of these terms. Differences in response probabilities (tolerance distributions) between the two groups defined for a particular variable may be indexed by differences in the constant term, the coefficient of noise on both. A useful measure of differences in responsiveness which takes into account both these sources of variation simultaneously is the "P-Effective Noise Level", denoted EN_P. This is the level of noise requisite to produce a response probability of P in a fitted model of response. Thus for the univariate logit model

$$EN_P = -\left(\frac{\alpha + \log(\frac{1}{P}-1)}{\beta}\right) \tag{3.4}$$

For example the .5-Effective Noise Level, $EN_{.5}$ is given by

$$\begin{aligned} EN_{.5} &= -\left(\frac{\alpha + \log(\frac{1}{.5}-1)}{\beta}\right) \\ &= -\frac{\alpha}{\beta} \end{aligned} \tag{3.5}$$

In fact, $(-\alpha/\beta)$ is the mean and median of a logistic distribution with parameters α and β, and the $EN_{.5}$ will be commonly referred to in the analysis as the "median effective noise". Clearly it only makes sense to calculate EN_P for models containing noise level as the single explanatory variable. In biometrics the EN_P is often interpreted as the noise level at which P·100% of the *population* become annoyed (Finney, 1971, 18). Note however that this extrapolation involves an implicit assumption of population homogeneity which is only satisfied when the type of stratification practiced here is undertaken. For a given stratum of the population the EN_p can be regarded as the noise which causes response in P·100% of the population of all individuals who possess the attributes which define that group.

A measure of the standardized difference between EN_p values derived from different samples is described in the next section, along with test statistics for hypotheses relating to model parameters. Measures of overall goodness of fit and associated test statistics are also described.

3.3 Test Statistics and Goodness of Fit Measures

In order to assess the overall explanatory power of the model developed in Chapter 4 we need some measure of overall fit to the calibration data set. The goodness of fit measure most commonly used with logit models is the likelihood ratio index (Tardiff, 1976). In its most general form this is defined as

$$\rho^2 = 1 - \frac{L^*(\underline{\underline{\beta}})}{L^*(H)}$$

$$= \frac{L^*(\underline{\underline{\beta}}) - L^*(H)}{-L^*(H)} \qquad (3.6)$$

where: $L^*(\underline{\underline{\beta}})$ = logarithm of the likelihood of the model with maximum likely parametervector $\underline{\underline{\beta}} = (\alpha, \beta_0, \beta_1, \ldots, \beta_k)$

$L^*(H)$ = log likelihood of some hypothesized base model.

Like the R^2 statistics used in regression analysis this index varies between 0 and 1. It provides a measure of the improvement in fit (or explanatory power) achieved by the fitted model over a base model which gives some assumed-minimum explanation (the true minimum of the log likelihood is of course undefined). The choice of a base model is clearly crucial. To date two specifications have been used.

The first of these is a model with all parameters set equal to zero ($L^*(H) = L^*(\underline{\underline{0}})$). This yields equal probabilities of responding and not responding for each individual since

$$P_i = [1 + e^{-0}]^{-1} = .5 = 1 - P_i \tag{3.7}$$

The justification for this base model is that when no information about the determinants of response is available a reasonable procedure is to assign $P_i = .5$ for all individuals. In an information-theoretic sense $-L^*(H) = -L^*(\underline{\underline{0}}) = -\sum_i \log .5$ corresponds to maximum entropy or uncertainty.

The second base model to be considered was first suggested by Tardiff who claimed that the conventional ρ^2 statistic based upon $L^*(\underline{\underline{0}})$ overestimated the fit of models by allowing the constant, α, to contribute to explanation. Tardiff's argument is that rather than using a base model with *all* coefficients equal to zero, fit should be considered in relation to a model with all coefficients zero *except the constant term* (denoted by $L^*(H) = L^*(C)$). This provides a measure of the explanatory power of the independent variables alone. The constant for the base model is chosen so as to maximize the likelihood of a model with a constant term only. Tardiff shows that for a binary logit model this implies

$$\alpha = \log\left(\frac{R}{N - R}\right) \tag{3.8}$$

where: R = total number responding in sample of size N.

Consequently,

$$P_i = [1 + e^{-\alpha}]^{-1} = \frac{R}{N}$$

and (3.9)

$$(1 - P_i) = 1 - \frac{R}{N} = \frac{N - R}{N}$$

so that the expected numbers responding/not responding agree with those observed. Since the base model now contains more information the likelihood ratio index calculated in this case, denoted ρ_c^2, will be less than or equal to ρ^2, with equality if and only if $\frac{R}{N} = .5$.

The selection of a measure of fit from these two depends upon how one views the role of the constant term in the model. It is important to note that Tardiff was concerned with qualitative *choice* models in which there is no theoretical justification for a constant. In choice-based travel demand models the constant is interpreted as indexing the effect of unobserved or unobservable variables on behaviour; but choice theory usually assumes, implicitly or explicitly, that all variables relevant to a choice can (*theoretically*) be measured. Certainly there is no place for a constant in the so-called abstract mode approach to travel demand modelling (Quandt and Baumol, 1966) which is based upon this assumption. In this context therefore the desire to exclude the contribution of the constant to a goodness of fit measure is understandable.

The same is not true however for the stimulus response models developed in Chapter 2. Here the fundamental theoretical concept on which the explanation of response hinges is that of a distribution of noise thresholds throughout the population. As the location parameter in this distribution, α contributes valid information about the nature of the response process, serving in some respects to characterize different stimuli (Goldstein, 1964). The role of the constant and its importance in this context are illustrated in the definition and description of the EN_p introduced above. One may therefore argue that the explanation due to the constant ought to be measured in any index of overall fit.

Clearly it is difficult, if not impossible, to decide which of these measures of fit is the "right" one. Rather than choosing one of them, both are presented in the present study, each to be interpreted according to the rationale which underlies it. Thus ρ^2 is a measure of the explanatory power of the entire vector of coefficients; ρ_c^2 of that of the independent variables only. Since the constant is largely determined by the sample split or aggregate frequencies of responders and non-responders, the latter measure can be seen as a measure of the information provided by the model over and above that which we obtain simply by knowing the aggregate response frequencies.

Associated with each of the two measures of fit are statistics which test the null hypothesis that (i) all of the t model coefficients are equal to zero and (ii) all of the coefficients except the constant are zero. These statistics are, respectively

$$\chi_t^2 = 2[L^*(\underline{\underline{\beta}}) - L^*(\underline{\underline{0}})] \quad (3.10)$$

$$\chi_{t-1}^2 = 2[L^*(\underline{\underline{\beta}}) - L^*(C)] \quad (3.11)$$

Both are asymptotically chi-square distributed, the former with t, the latter with t-1 degrees of freedom (Cox, 1970).

In order to operationalize the algorithm presented in Figure 3.1 we also need statistics with which to assess the significance of individual variable coefficients and of differences in fit between models containing different numbers of variables. A chi-square statistic is again used to test the null hypothesis that a model with t coefficients offers no improvement over one with s < t coefficients. This statistic is given by

$$\chi_{t-s}^2 = 2[L^*(\underline{\underline{\beta}}^t) - L^*(\underline{\underline{\beta}}^s)] \quad (3.12)$$

where $L^*(\underline{\underline{\beta}}^t)$ = maximum log likelihood of the model with t coefficients. In this case the number of degrees of freedom is (t-s).

The significance of individual coefficients is evaluated using a t-statistic. The null hypothesis is that a particular coefficient estimate, $\hat{\beta}_k$, does not differ significantly from zero. The t-statistic is defined as:

$$t = \frac{\hat{\beta}_k}{(v_{kk})^{1/2}} \quad (3.13)$$

where v_{kk} = the kth diagonal element of the inverse of the negative of the

Hessian matrix of second partial derivatives of the likelihood functions. This definition makes use of the fact that the *asymptotic* sampling distribution of the maximum likelihood estimator $\underline{\underline{\hat{\beta}}}$ is normal with expected value $\underline{\underline{\beta}}$, and variance matrix

$$\underline{\underline{V}} = \underline{\underline{W}}^{-1} \tag{3.14}$$

where

$$\omega_{ij} = -E\left(\frac{\partial^2 \log L}{\partial\beta_i \; \partial\beta_j}\bigg|\; \beta_i, \beta_j\right)$$

$$= -\left(\frac{\partial^2 \log L}{\partial\beta_i \; \partial\beta_j}\bigg|\; \beta_i, \beta_j\right)$$

(Domencich and McFadden, 1975, 111-112). A consistent estimate of ω_{ij} is given by

$$\hat{\omega}_{ij} = -\left(\frac{\partial^2 \log L}{\partial\beta_i \; \partial\beta_j}\bigg|\; \hat{\beta}_i, \hat{\beta}_j\right) \tag{3.15}$$

which is the estimate used to calculate the t-statistics.

It would be useful to have a similar statistic to measure the significance of differences in EN_p values estimated for different subgroups of the sample. By analogy to the t-statistic we can define a measure - call it d - given by:

$$d = \frac{\hat{EN}_p(A) - \hat{EN}_p(B)}{(\sigma_A^2 + \sigma_B^2)^{1/2}} \tag{3.16}$$

where $\hat{EN}_p(A)$ = estimated EN_p of group A

σ_B^2 = variance of $\hat{EN}_p(B)$

The estimated EN_p is found simply by evaluating equation (3.4) at $\hat{\alpha}$, $\hat{\beta}$: the variance of this estimate is derived in Appendix 2. Like the t-statistic for the difference in two parameters, d measures the difference between the $E\hat{N}_p$'s of two subgroups in units of the standard error of this difference (the latter derived on the assumption that the covariance of $E\hat{N}_p(A)$ and $E\hat{N}_p(B)$ is zero).

The problem is however that d does not have a convenient sampling distribution. In Appendix 2 it is shown that the probability density function for $E\hat{N}_p$ is composed of a term which is something like a Cauchy density function and a term which resembles a normal density function. Because of time constraints it was not possible to go further and derive the distribution of d but this is likely to be extremely complex. It certainly does not correspond to any of the well-known and documented sampling distributions. Moreover the Cauchy term in the density function raises a further problem. In the special case that the distributions of both $\hat{\alpha}$ and $\hat{\beta}$ are located at the origin (which is likely to occur when the fit of the model is poor) the "normal" term disappears so that the sampling distribution of the $E\hat{N}_p$ is a modified Cauchy. But the moments of the Cauchy are undefined which means that significance tests cannot be conducted on parameters with this distribution. This plus the fact that the sampling distribution of the d-statistic is unknown means that significance levels cannot be calculated for this statistic. However the d itself can still be employed in a purely descriptive capacity where comparisons are drawn between group $E\hat{N}_p$ values.

Appendix 3 contains details of the numeric method used to solve the maximum likelihood equations for the parameter vector $\underline{\underline{\beta}}$. A copy of the computer estimation program, including a subroutine to calculate the statistics described above, is given in Appendix 4.

3.4 Summary

This chapter has been concerned with operationalizing the model of response described in Chapter 2.

The day-night equivalent sound level, L_{dn}, is chosen as the noise metric, while attitudes towards both overall neighbourhood noise and main road traffic noise are selected as the response measures. In all, four response variables are defined using two different cut-off points in the response scales for each of the two noise sources. The battery of intervening variables comprises a wide array of socio-economic, demographic, psychological and situational characteristics of individuals.

The explanatory variables may either be incorporated directly into multivariate response equations, or used as the basis of a sample stratification designed to permit differences in response across subgroups to be examined. It was decided to use both methods of examining the intervening variables in the present study, giving rise to a two-stage analysis. In stage 1 variables will be entered in turn into multivariate response models and retained if a number of pre-determined tests are satisfied. Variables with ordinal and nominal properties have to be transformed into dummy variables before being entered. The variables which are retained will then be further examined in stage 2 for differences in the responsiveness of groups characterized by different levels of these attributes.

The measure of responsiveness to be used is the noise level associated with given probabilities, P, of response in the subgroup in question. This measure is denoted with "P-Effective Noise Level" or EN_p. Unfortunately the sampling distribution of the EN_p estimate is unknown, precluding significance tests of differences in this response measure across groups. However all of the other statistics used in the analysis can be subjected to significance tests since they all have well-known sampling distributions.

CHAPTER 4

MULTIVARIATE MODELS OF RESPONSE

This chapter investigates the effect of various non-acoustical variables on attitudes towards both overall neighbourhood noise and main road traffic noise. In all, four attitudinal variables are examined, as described in Chapter 3, and the structure of the chapter is based on the presentation of the results for each of these in turn. The analysis of the first dependent variable, which measures whether or not the respondent was at all disturbed by neighbourhood noise levels, is reported in some detail in order to illustrate the procedure by which the best model specification for a particular response variable is derived. The treatment of the other dependent variables will be briefer and concentrated more on the best-fit model specification. The chapter concludes with a summary and discussion of the results.

Note that the abbreviated forms of the names of the four response variables (as given in Table 3.1) will be used frequently throughout the discussion. The numbers in parentheses after these abbreviations refer to the various cut-off points used to define the response variables. Thus "1-5/6-9" indicates a neutral cut-off; "0-7/8-9" an extreme cut-off in which persons who did not respond to the question are classified as "not severly disturbed".

4.1 Attitude Towards Overall Neighbourhood Noise - Any Disturbance

Analysis of *NHDATT (1-5/6-9)* began with the estimation of a model which contained noise level (L_{dn}) as the single explanatory variable. The results (Table 4.1) indicate that though noise level is positively (and significantly) related to the incidence of disturbance the overall model fit is poor as indexed by a ρ_c^2 of .06. As further evidence of the failure of noise level alone to explain individual response, this finding is in accord with those of Langdon and Griffiths (1968) and Langdon (1976), and is an appropriate point of departure for the present analysis, which investigates the remaining unexplained variation.

The first intervening variable entered in the predictive equation was sensitivity. This was converted into a dummy variable with value 1 for respondents who described themselves as more sensitive than average and 0 for all others. This particular cut-off point was chosen arbitrarily: for the rest of the intervening variables however, break points were chosen as described in Chapter 3. The result of including sensitivity is almost a doubling of the level of explanation (Table 4.2), a similar result to that observed by Langdon (1976). The coefficient of sensitivity is positive and highly significant[1], indicating that more highly sensitive individuals are, *ceteris paribus*, more likely to be disturbed. The χ^2-statistic for improvement in fit over the base model (model 1.1) is also significant,

---

1 The critical values set for significance levels were .01 in the case of the χ^2-statistic for improvement in fit, and .05 for the t-statistic. The t-tests were one- or two-tailed according to whether there was a prior hypothesis as to the sign of the coefficient: for example the significance levels for L_{dn} are in all cases those derived from one-tail of the t-distribution. If the t-statistic for a variable added to a model was not significant the χ^2-statistic was not computed, that variable being automatically eliminated from further consideration.

Table 4.1

NHDATT (1-5/6-9) model 1.1

Variable	Coefficient	t
CONSTANT	-6.16	11.73
L_{dn}	0.08	10.57

ρ^2	.13	χ_2^2	308.41
ρ_c^2	.06	χ_1^2	121.81
N†	1700		

† N is the sample size.

Table 4.2

NHDATT (1-5/6-9) model 1.2

Variable	Coefficient	t
CONSTANT	-7.46	9.89
L_{dn}	0.10	8.72
SENSE (greater than average)	0.89	4.68

ρ^2	.21	χ^2_3	273.43
ρ^2_c	.11	χ^2_2	108.50
N	821	†χ^2_1	21.65

† This is the statistic to test the significance of the improvement in fit over the previous best model - in this case over model 1.1.

and so sensitivity is retained as an explanatory variable and model specification 1.2 becomes the new base model.

It should be noted that the sample size is more than halved by inclusion of sensitivity in the model. This is due to the fact that sensitivity was only measured in the 1975 survey and it is from this that the 821 observations in model 2.1 are drawn. Because of the importance of sensitivity throughout, the analysis is effectively limited to the 1975 data set. The sample size will decrease further (though much less drastically) with inclusion of other variables on which there are observations with missing data.

After sensitivity a number of other intervening variables were tested, including *INCOME*, number of hours spent at home during the week (*WKHME*) and hours at home during the weekend (*WENDHM*). In no case were the coefficients of these variables significantly different from zero at the desired confidence level. The next variable to have a significant coefficient was in fact the individual's rating of the noise at his place of work (*RATWORK*). The cutting point on the rating scale for this variable was such as to discriminate between respondents who rated their work-place *quieter* than average and all other respondents. The estimation results (Table 4.3) reveal that individuals whose work-places are relatively quiet are less likely to report disturbance than respondents who are similar with respect to the other variables in the model but who rate their work-place as noisier.

This finding is interesting and informative since it is difficult to decide *a priori* the likely effect of work-place noise on attitude towards neighbourhood noise. Plausible arguments could be made to support

Table 4.3

NHDATT (1-5/6-9) model 1.3

Variable	Coefficient	t
CONSTANT	-6.89	8.66
L_{dn}	0.09	8.01
SENSE	0.94	4.68
RATWORK (quieter than average)	-1.10	5.97

ρ^2	.25	χ^2_4	274.821
ρ^2_c	.15	χ^2_3	145.87
N	801	χ^2_1	38.58

hypotheses of either a positive effect (as observed) or a negative effect (i.e. quiet work-places associated with disturbance by neighbourhood noise). What the present result suggests is that disturbance is additive in the sense that though it may be attributed by the individual to a particular source, disturbance may in fact reflect the combined effect of a number of sources. In other words the individual's tolerance for noise experienced in the vicinity of the home is not independent of noise experienced elsewhere. Since both its coefficient and contribution to fit were significant, *RATWORK* was included in the model for *NHDATT (1-5/6-9)*. The result of its inclusion in terms of the overall fit is to raise ρ_c^2 by 1/3 of its previous value.

The next variable to be tested, relaxation interruption, (*RLXINT*), simply recorded whether or not an individual had suffered interruption of relaxation by the noise source in question. As expected, individuals who *were* interrupted are more likely to express disturbance (Table 4.4). The magnitude of the coefficient of *RLXINT*, its associated t-statistic, and the improvement in model fit are encouraging, and indicate a strong relationship between disturbance and activity interruption of this kind.

Variables measuring the interference of noise with work, conversation, eating, television viewing and telephone conversations were tested after *RLXINT*, but none of them yielded a significant improvement in fit over model 1.4, and only work interruption had a significant coefficient. It is likely however that this result is due to intercorrelation among the activity interference variables. Although this was not tested for directly, it was found that when each of the activity interference measures listed above was entered by itself into model 3.1 it yielded a significant

Table 4.4

NHDATT (1-5/6-9) model 1.4

Variable	Coefficient	t
CONSTANT	-5.99	7.27
L_{dn}	0.07	6.05
SENSE	0.82	3.83
RATWORK	-1.03	5.35
RLXINT	1.51	7.66

ρ^2	.30	χ^2_5	334.48
ρ^2_c	.21	χ^2_4	205:53
N	801	χ^2_1	59.66

coefficient and improvement in fit. None of these performed quite as well as *RLXINT* however and so it was decided to retain this particular variable as the best representation of what appears to be a single underlying dimension of variation. These findings help to illustrate the fact that the intermediate results described here are not independent of the order of incorporation of the intervening variables into the model specifications (though the final 'best' model specifications are). In the present study this order was in part arbitrary, in part based upon intuition as to what variables were likely to contribute most to explanation of response - those thought to be most powerful being entered first.

Two demographic variables sex and age were tested next: both gave very poor results.

At this point it was decided to explore alternative specifications of the sensitivity variable, the cut-off point used up to and including model 4.1 having been chosen arbitrarily. It was found that by placing the break point immediately *below* the neutral or medium point on the scale a better explanation of disturbance was achieved (Table 4.5), and so this new specification was adopted for all subsequent model estimations. The gain in explanatory power is indexed by the slight increase in the absolute magnitude of the variable coefficient: the sign of the coefficient changes to negative of course since the variable is now specific to the *less* sensitive part of the sample. Note that since no new variables are being entered into the model it was not necessary (nor appropriate) to test the *significance* of the improvement in fit.

Following this respecification the remaining intervening variables - length of residence, shielding, level of education, tenure and sleep

Table 4.5

NHDATT (1-5/6-9) model 1.5

Variable	Coefficient	t
CONSTANT	-5.49	6.66
L_{dn}	0.08	6.07
SENSE (less than average)	-0.89	4.77
RATWORK	-1.00	5.16
RLXINT	1.41	7.09

ρ^2	.31	χ^2_5	343.41
ρ^2_c	.22	χ^2_4	214.46
N	801		

interruption - were tested. Of these only sleep interruption was retained, to give model specification 1.6 (Table 4.6). It is worth noting however that the coefficients of both length of residence (positively related to disturbance) and tenure (owner occupiers more likely to be disturbed) were significant, while their χ^2-statistics only just failed to reach significance at the predetermined level. Sleep interruption on the other hand barely satisfied the requirements for entry, which is somewhat surprising in view of the importance placed upon it elsewhere (e.g. Langdon and Buller, 1977). The explanation of this result may be due in part to intercorrelation between *SLPINT* and L_{dn}. Clearly, some degree of correlation is to be expected from the very definition of L_{dn} (see Chapter 3, Section 3.1). Further support for the hypothesis comes from another study based upon the present data set (Taylor, Gertler and Hall, 1978) which found a significant relationship between disturbance and sleep interruption: significantly, no measure of noise level was included as a predictor.

A final attempt to improve the fit of the model consisted in constructing a composite variable called "speech interference" which had value 1 if verbal communication in general was interfered with, during either person to person or telephone conversations, or listening to television. The variable failed however to support a significant coefficient. Model 1.6 is therefore the best model obtainable for this particular response variable.

4.2 Attitude Towards Overall Neighbourhood Noise - Severe Disturbance

The first model calibrated for *NHDATT (1-7/8-9)* was the same as that which had given the best fit for the same dependent variable defined

Table 4.6

NHDATT (1-5/6-9) model 1.6

Variable	Coefficient	t
CONSTANT	-5.36	6.45
L_{dn}	0.07	5.72
SENSE	-0.82	4.38
RATWORK	-1.02	5.20
RLXINT	1.22	5.81
SLPINT	0.67	2.94

ρ^2	.32	χ^2_6	351.98
ρ^2_c	.23	χ^2_5	223.02
N	801	χ^2_1	8.57

according to the neutral cut-off point, the only change being in the specification of *RATWORK* (see Table 4.7). Estimation of this model resulted in non-significant coefficients for both *RLXINT* and *SLPINT*. Both were dropped from the vector of explanatory variables and at the same time *CONINT* was added to give model 2.2 (Table 4.8). This respecification resulted in a highly significant coefficient of *CONINT* and an improvement in overall fit. This leads to the interesting conclusion that whilst interruption of sleep or relaxation produces mild or moderate disturbance, conversation interruption causes more extreme disturbance. This justifies the importance attached to the latter variable in for example the USEPA "Levels Document" (USEPA, 1974), and corroborates the findings of numerous other studies.

Another important point brought out by a comparison of this model with model 1.6 is the substantial improvement in fit obtained in the present case. Evaluation of the magnitude of this improvement depends though on the measure of fit on which the comparison is based, for the two indices diverge quite widely in model 2.2. The reason for this divergence is that the sample split is now quite extreme with only 7% of the respondents describing themselves as highly disturbed. A model with a constant alone will therefore give quite a good explanation of response and so ρ_c^2 is low relative to ρ^2. Yet the former measure shows a significant increase over the corresponding value in model 1.6, indicating that even allowing for differences in the sample split, it is easier to predict response as defined according to the more extreme break point. This lends support to arguments in favour of the use of "percentage highly annoyed" in setting noise standards (Schultz, 1977; Rylander, Sorensen and Kajland,

Table 4.7

NHDATT (1-7/8-9) model 2.1

Variable	Coefficient	t
CONSTANT	-8.60	6.09
L_{dn}	0.10	5.10
SENSE (less than average)	-1.28	3.22
RATWORK (noisier than average)	1.78	5.28
RLXINT	0.62	1.89
SLPINT	0.33	0.99

ρ^2	.74	χ^2_6	816.86
ρ^2_c	.29	χ^2_5	122.68
N	801		

Table 4.8

NHDATT (1-7/8-9) model 2.2

Variable	Coefficient	t
CONSTANT	-7.70	5.35
L_{dn}	0.09	4.42
SENSE	-1.39	3.52
RATWORK	1.82	5.33
CONINT	1.24	3.47

ρ^2	.74	χ^2_5	822.621
ρ^2_c	.31	χ^2_4	128.443
N	801		

1976).

None of the other intervening variables contributed significantly to the explanation of *NHDATT (1-7/8-9)*, and so model specification 2.2 represents the best fit model for this response variable.

4.3 Attitude to Main Road Traffic Noise - Any Disturbance

Turning attention now to response to main road traffic noise, two characteristics of the data set which have a bearing on the analysis need to be noted. First, individuals were asked to rate their disturbance due to noise from main road traffic *only if* they indicated that they had noticed it. Many individuals did not and so no rating exists in these cases. Rather than discarding these observations however it was felt safe to assume that if a respondent had not even noticed the noise he was not disturbed by it. Such individuals are therefore classified as non-disturbed observations on both the dichotomized dependent variables defined for this noise source. Secondly, none of the activity interference variables could be included in model specifications for either of these response variables. This was because respondents who did not report any disturbance by main road traffic noise were not questioned further about activity interference arising from this noise source. It does not seem justified to assume that these individuals would not have reported activity interference if asked: obviously an individual may suffer, say, interruption of conversation without necessarily being annoyed or disturbed. There is therefore no alternative other than to exclude these variables from further analyses, and this should be borne in mind when drawing comparisons with the model specifications for the two *NHDATT* variables.

The results from the first model evaluated for *MNATT (0-5/6-9)* (Table 4.9) indicate that noise level alone (plus a constant) gives its best performance so far, the ρ_c^2 index for model 3.1 being almost twice that for model 1.1 for example. The improvement in fit and the increased magnitude of the coefficient of L_{dn} suggest that response to main road traffic noise is better predicted by the noise measure alone than is response to overall neighbourhood noise. The explanation of this finding may lie in the fact that the latter variable probably taps a much wider range of attitudes and feelings, with the result that non-acoustical variables could be expected to play a larger role in its determination.

Adding sensitivity to the model specification results in a substantial increase in fit (Table 4.10), providing further evidence of the importance of this variable in explaining response.

The best model specification for *MNATT (0-5/6-9)* contains, in addition to sensitivity, only one other intervening variable, namely level of education (Table 4.11). The positive coefficient of *LEVED* indicates that respondents whose level of educational attainment includes at least high school graduation are more likely to register disturbance than those with less formal education, (at most some high school). Initially *INCOME* too had entered into the model specification but lost its explanatory power when *LEVED* was added, due to the correlation between the two. This is an intriguing result in view of the fact that it is commonly assumed (e.g., Borsky, 1970) that socio-economic indicators bear no relation to disturbance (though, as discussed earlier they may, and usually do, influence such indirect measures of disturbance as complaint activity). The present finding challenges this assumption and suggests that the spatial distribution

Table 4.9

MNATT (0-5/6-9) model 3.1

Variable	Coefficient	t
CONSTANT	-8.95	15.39
L_{dn}	0.12	14.41

ρ^2	.19	χ^2_2	445.29
ρ^2_c	.11	χ^2_1	250.12
N	1715		

Table 4.10

MNATT (0-5/6-9) model 3.2

Variable	Coefficient	t
CONSTANT	-9.24	10.16
L_{dn}	0.12	9.31
SENSE (less than average)	-1.32	6.12

ρ^2	.42	X_3^2	481.76
ρ_c^2	.18	X_2^2	148.68
N	832	X_1^2	42.2

Table 4.11

MNATT (0-5/6-9) model 3.3

Variable	Coefficient	t
CONSTANT	-9.21	9.94
L_{dn}	0.13	9.32
SENSE	-1.26	5.79
LEVED (at least high school graduation)	0.73	3.48

ρ^2	.43	χ^2_4	492.19
ρ^2_c	.19	χ^2_3	159.13
N	829	χ^2_1	12.66

of response may be in part determined by the distribution of variables such as income or education. Other things, including L_{dn}, being equal, response appears to be greater in areas of higher socio-economic status. One explanation of this result which seems plausible is that those with lower levels of education (or lower socio-economic status in general) tend to have a lower aspiration level with respect to their acoustic environment and, as a result, are less likely to suffer disturbance or stress due to noise (c.f. for example Brummell, 1977, 116-120). If this is the case then it would obviously be wrong, from a social-justice point of view, to use the present finding to support a socially discriminating noise-control policy.

4.4 Attitude Towards Main Road Traffic Noise - Severe Disturbance

The combination of variables which best explains "considerable" or "extreme" disturbance due to main road traffic noise consists of only the noise level and sensitivity (Table 4.12). Both of these variables are highly significant as is the overall model fit. The improvement in fit over model 3.3 is further evidence of the fact that response defined according to the more extreme cut-off point is easier to predict. It is interesting to note, in connection with the discussion above that *LEVED* disappears from the model specification. This lends support to Schultz's arguments in favour of the more extreme break point on the grounds that "the effects of non-acoustical variables are reduced" (Schultz, 1977, 6). However it should be noted that sensitivity once again contributes significantly to model fit. In fact addition of this variable raised the ρ_c^2 statistic from .17 for a model with just a constant and L_{dn} to .26 in model 4.1.

Table 4.12

MNATT (0-7/8-9) model 4.1

Variable	Coefficient	t
CONSTANT	-12.0	8.43
L_{dn}	0.15	7.30
SENSE (less than average)	-2.49	4.65

ρ^2	.74	χ^2_3	858.98
ρ^2_c	.26	χ^2_2	105.37
N	832		

4.5 Summary and Discussion

This chapter has developed predictive equations to explain response to noise as defined according to 4 different response variables. Focussing first upon differences in the results for the two different noise sources we are led to conclude that attitude towards main road traffic noise is in general easier to predict and better explained by noise level alone than is attitude to overall neighbourhood noise. The first part of this conclusion is based on the fact that although the goodness of fit statistics are better for the two *NHDATT* models, this is due to the inclusion of activity interference variables in these models, which, because of the method in which the data were collected could not be used to predict attitude towards main road traffic noise. Compare for example models 1.3 and 3.3, neither of which contain activity interference variables. The claim that noise level alone predicts *MNATT* better than *NHDATT* is borne out by the relatively larger coefficients of L_{dn} in models 3.3 and 4.1 as compared to models 1.6 and 2.2; and also by the better fit of model 3.1 compared to model 1.1. These results are to be expected since attitude towards overall neighbourhood noise is likely to be influenced by a much wider range of non-acoustical variables than the more specific attitude to main road traffic noise. Langdon (1976) found for example that overall satisfaction with the neighbourhood was important in this respect. In the present case it is informative to note that *RATWORK* is one of the most important variables in the predictive equations for *NHDATT*, but fails to appear in the *MNATT* model specifications. This is probably due to the fact that there is an obvious comparison to be made between noise levels at home and at work, a comparison which was certainly implicit in the questionnaire and which leads to the fairly strong relationship observed between *NHDATT* and *RATWORK*. When the individual is asked to focus specifically on noise from main road traffic however the comparison is less obvious and the relationship vanishes.

Looking next at the differences in results for the two specifications of response it is clear that severe disturbance is easier to predict than any disturbance. This is probably due to the fact that, as Schultz (1977) points out, response measures based on neutral break points are likely to be more subject to random variation than those defined by more extreme cut-off points. Comparison of the magnitude of the coefficients of L_{dn} in models 1.6 and 2.2 with the magnitude of those in models 3.3 and 4.1 prompts the conclusion that noise level alone is a better predictor of severe disturbance than it is of any disturbance.

As regards the performance of individual explanatory variables, the results indicate that attention can be confined to a relatively small group of intervening variables: these are sensitivity, rating of work-place noise, level of education, and interruption of relaxation, sleeping and conversation. By virtue of its appearance in all four of the best model specifications, sensitivity is probably the most powerful of these explanatory variables. It is essentially a psychological variable, and though it is encouraging to be able to reduce some of the unexplained variations in individual response, one can question the utility or information value of this result from a predictive point of view on the grounds that in any practical situation where predictions of response are required from a small amount of prior information about the exposed population, sensitivity will be unknown and unpredictable.

In general the ρ_c^2 statistics for the 4 best model specifications are not very impressive, ranging from .19 for *MNATT (0-5/6-9)* to .31 for *NHDATT (1-7/8-9)*. In view of these low levels of explanation, the major conclusion to be drawn from this part of the analysis must be that variation in individual response to noise remains largely unexplained. However it is felt that the failure to achieve a more complete explanation of response is due not so much to any misspecification of the form of model chosen for analysis, but to gaps in the data on individual characteristics. For example, the goodness of fit statistics for the two main road traffic noise response variables would probably have been considerably higher had activity interruption data for this noise source been available in usable form. Similarly, some measure of the individual's overall satisfaction with his neighbourhood, if available, may have helped reduce the unexplained variation in attitudes towards neighbourhood noise. Moreover, the levels of explanation actually achieved compare favourably with those reported elsewhere for individual response to noise. For example the highest R^2 obtained by Langdon (1976, 258, Table 12) was only .2, which is just above the *lowest* ρ_c^2 obtained in the present study. This improvement, considered together with the much stronger theoretical and statistical basis of the present modelling approach admits some optimism.

CHAPTER 5

INTER-GROUP COMPARISON OF RESPONSE MEASURES

The objective in this chapter is to investigate differences in response to noise across various groups of the sample. The analytic design has already been described in Chapter 3, so no further comments are necessary here except to note that the inter-group comparisons are based on the $EN_{.5}$ (see equation 3.5), i.e., the level of noise required to produce the response of interest ("any disturbance", "severe disturbance") in 50% of the individuals in the group. This particular level of the EN_p statistic is chosen for two reasons: firstly because of the obvious planning significance of the 50% response level; and secondly because as a measure of central tendency the $EN_{.5}$ is less subject to sampling variation than are more extreme measures such as $EN_{.1}$ and $EN_{.9}$ (Finney, 1971, 18-19).

Following the format of the previous chapter, the first four sections present the results for each of the dependent variables in turn. The final section summarizes the more important findings and compares the results to those obtained by other researchers. This comparison helps to illustrate the biases (errors) that may be introduced to the setting of noise standards by failing to take into account the effect of intervening, non-acoustical variables.

5.1 Differences in Median Effective Noise for Any Disturbance by Overall Neighbourhood Noise

The results of the multivariate model estimations reported in Chapter 4 suggest that there are four intervening variables which affect the incidence of disturbance due to overall neighbourhood noise levels. These are sensitivity, rating of work-place noise, and interruption (if any) of relaxation and sleeping.

Analysis began by looking at differences in the $EN_{.5}$ between individuals who are less sensitive[1] than average and all other respondents, whilst holding the remaining 3 intervening variables constant. Though the results (Table 5.1) appear to indicate a large difference in the median effective noise for these two groups, the more striking aspect of the data is the lack of fit of the response equation for the less sensitive group. The adjusted chi-square index of fit is not significant at the required confidence level (.01), whilst the coefficient of L_{dn} barely achieves significance. Clearly, disturbance by overall neighbourhood noise within this group is only weakly related to the noise level itself. Because the coefficient of noise is close to zero the standard deviation of the $EN_{.5}$ estimate is very large, (see equation A2.2), so that little confidence can be placed in the value of 95 dBA obtained. The goodness of fit statistics for the more sensitive group are better, and the standard error of the $EN_{.5}$ estimate much lower; but because of the large variance of the $EN_{.5}$ for the less sensitive group, the standardized difference, d,

---

1 Throughout this part of the analysis the cut-off points used are those on which optimum model specifications were based in Chapter 4.

Table 5.1

Differences in NHDATT (1-5/6-9) according to sensitivity:

RATWORK, RLXINT and SLPINT held constant

	Less Sensitive Group	More Sensitive Group
N	160	138
$\log(\hat{P}/1-\hat{P})$†	$-5.04 + 0.05\ L_{dn}$	$-6.67 + 0.09\ L_{dn}$
t	2.73 1.90	3.71 3.34
ρ^2	.36	.14
ρ_c^2	.02	.07
χ^2	80.15	26.35
χ_c^2	3.59	13.36
$\hat{E}N_{.5}$	95.11	70.52
σ	15.86	2.77

d = 1.53

† Since $\hat{P} = (1 + e^{-\hat{\alpha}-\hat{\beta}L_{dn}})^{-1}$

$$\frac{1}{\hat{P}} - 1 = \frac{1-\hat{P}}{\hat{P}} = e^{-(\hat{\alpha}+\hat{\beta}L_{dn})}$$

so, $$\log\left(\frac{\hat{P}}{1-\hat{P}}\right) = \hat{\alpha} + \hat{\beta}L_{dn}$$

is relatively small. As a result it is difficult to conclude anything definite about differences in response across these two groups.

There are at least two possible explanations of this finding. On the one hand it may be that there is insufficient variation in noise levels within the less sensitive group to support a strong relationship between L_{dn} and response. This would occur if the noise level and sensitivity were correlated. It has, for example, been suggested that less sensitive individuals may tend to live at relatively more noisy sites than more sensitive individuals. Time did not permit of a full test of this hypothesis but some support is lent it by the fact that whereas the variance in noise levels among the more sensitive group (which includes average individuals) was 57.65 dBA^2, it was only 36.48 dBA^2 for the less sensitive group.

An alternative explanation of the results is that the less sensitive group may contain a large number of what Schultz (1977) has termed "imperturbables"; that is, individuals who remain undisturbed by any level of noise within the range commonly experienced at residential sites. Weight is added to this explanation when it is considered that for the analysis of sensitivity both activity interruption variables were held constant at their "base" values; i.e., these groups represent individuals who reported no interruption of either sleep or relaxation by noise. Further support comes from the fact that of the 160 respondents in this group 133, (83.1%), were not at all disturbed.

Quite probably the present results reflect both of these factors, i.e., lack of variation in the noise levels within the group and the presence of "imperturbable" individuals. In an attempt to both increase the variability of L_{dn} and add individuals who were likely to have been disturbed, it was decided to relax control of both the activity interruption variables. The result (Table 5.2) is as anticipated, with the fit of the model for the less sensitive group, being improved sufficiently for L_{dn} to now support a significant coefficient. Again there exists a large difference in the median effective noise between the two groups, ranging from 83.27 dBA for the less sensitive group to 65.45 dBA for the more sensitive group. The d-statistic indicates that the difference between these two levels is more than 2½ times the standard error of the difference, prompting the conclusion that response to neighbourhood noise differs quite markedly between the more and less sensitive segments of the population.

The next variable to be examined was rating of work-place noise. Holding sensitivity, sleep interruption and relaxation interruption constant, separate models were calibrated for those individuals who reported quieter than average work-places and all other respondents (Table 5.3). As was the case for the less sensitive individuals, it was impossible to obtain a significant fit for either of the two *RATWORK* groups. Moreover, for those individuals with quieter work-places, the coefficient of noise itself failed to achieve significance. The $EN_{.5}$ statistic indicates very little, if any, difference between these groups, but this is again due to the poor performance of the noise measure, L_{dn}, in explaining response. If we look instead at the constants in the two response equations we are led to believe that there is in fact a considerable difference between the groups. These apparently contradictory findings are due to the fact that though there is no strong relationship between L_{dn} and response in either group - so that, in a trivial sense, the groups are similar in this respect - there are

Table 5.2

Differences in NHDATT (1-5/6-9) according to sensitivity:

RATWORK held constant

	Less Sensitive Group	More Sensitive Group
N	206	262
$\log(\hat{P}/1-\hat{P})$	$-5.11 + 0.06\ L_{dn}$	$-6.83 + 0.10\ L_{dn}$
t	3.56 2.85	5.44 5.47
ρ^2	.21	.10
ρ_c^2	.04	.10
X^2	61.15	37.06
X_c^2	8.34	36.92
$\hat{E}N_{.5}$	83.27	65.45
σ	6.50	1.27

d = 2.69

Table 5.3

Differences in NHDATT (1-5/6-9) according to rating of work-place noise: SENSE, RLXINT and SLPINT held constant

	Quieter Work-place	Noisier Work-place
N	143	160
$\log(\hat{P}/1-\hat{P})$	$-8.90 + 0.09\ L_{dn}$	$-5.04 + 0.05\ L_{dn}$
t	2.24 1.57	2.73 1.90
ρ^2	.70	.36
ρ_c^2	.04	.02
χ^2	139.05	80.16
χ_c^2	2.49	3.60
$\hat{E}N_{.5}$	94.10	95.11
σ	18.44	15.86

d = 0.04

substantial differences in the proportions disturbed, and it is these proportions which largely determine the constant terms.

Once again it was decided to relax control of the two activity interruption variables in order to introduce more variation in both L_{dn} and the response variable. This produced the desired result, (Table 5.4), namely significant fits for both groups and low variance estimates of the $EN_{.5}$. The data indicate that holding sensitivity constant, there is a large difference in the median effective noise according to whether or not the individual rates his work-place as quieter than average. For those who do enjoy relatively quiet places of employment the noise level at which a typical individual has .5 probability of being disturbed is almost 10 dBA higher than the corresponding mark for individuals with more noisy work-places: this difference in the $EN_{.5}$ is almost 3 times the standard error of the difference.

Before proceeding further it is necessary to ask What is the effect of allowing relaxation interruption and sleep interruption to vary when looking at differences in the $EN_{.5}$ levels across different *SENSE* and *RATWORK* groups? In particular, to what extent are the differences in response levels which emerge in Tables 5.2 and 5.4 due to coincident inter-group differences in the proportions of individuals who experienced relaxation and sleep interruption? A complete answer to this question would require a comprehensive analysis of the structure of the intercorrelations between the intervening and response variables, and would certainly carry us beyond the scope of the present study. On the basis of what data there are available however it is clear that there is covariation between activity interruption and both *SENSE* and *RATWORK* which may have a bearing on the

Table 5.4

Differences in NHDATT (1-5/6-9) according to rating of work-place noise: SENSE held constant

	Quieter Work-place	Noisier Work-place
N	166	262
$\log(\hat{P}/1-\hat{P})$	$-8.55 + 0.11\ L_{dn}$	$-6.83 + 0.10\ L_{dn}$
t	4.40 3.86	5.44 5.47
ρ^2	.30	.10
ρ_c^2	.09	.10
χ^2	68.49	37.06
χ_c^2	16.97	36.92
$\hat{E}N_{.5}$	75.26	65.45
σ	3.02	1.27

d = 3.00

interpretation of the results. For example, whereas only 17% of the less sensitive group reported disturbance of relaxation, the same statistic for the more sensitive group was 38%.

However, there are two arguments in favour of overlooking the effects of variation in activity interference and interpreting the results in Tables 5.2 and 5.4 as due to *SENSE* and *RATWORK* respectively. First, the highest intercorrelation (phi-coefficient) of relevance to this discussion was .20 for the relationship between sensitivity and relaxation interruption. This seems toosmall to be the source of a large aggregation bias. Secondly, and more importantly it is argued that activity interruption measures are not truly independent variables but rather functions of other variables such as sensitivity and work-place noise (c.f. Section 3.1). As a result it seems safe to attribute the variation in *both* median effective noises *and* percentages suffering activity interruption which emerges in, for example, Table 5.2, to the difference in sensitivity across the two groups. On the strength of these arguments the failure to adhere to "other things being equal" conditions in deriving and interpreting the results in Tables 5.4 and 5.4 is justified.

This now imposes some constraints, if not on the analysis of the activity interruption variables, then certainly on the interpretation of the results. Clearly it is desirable that sensitivity and work-place rating be controlled in treating the two interruption variables, in order to ascertain the extent to which *RLXINT* and *SLPINT* independently affect response.

Unfortunately, the results for both *RLXINT* (Table 5.5) and *SLPINT* (Table 5.6) reveal once again the inability of L_{dn} alone to explain

Table 5.5

Differences in NHDATT (1-5/6-9) according to relaxation interruption: SENSE, RATWORK and SLPINT held constant

	No Relaxation Interruption	Relaxation Interruption
N	138	54
$\log(\hat{P}/1-\hat{P})$	$-6.67 + 0.09\ L_{dn}$	$-1.89 + 0.04\ L_{dn}$
t	3.71 3.39	0.57 0.80
ρ^2	.14	.11
ρ_c^2	.07	.01
χ^2	26.35	8.26
χ_c^2	13.36	0.68
$\hat{E}N_{.5}$	70.53	48.53
σ	2.77	25.55

d = 0.86

Table 5.6

Differences in NHDATT (1-5/6-9) according to sleep interruption: SENSE, RATWORK and RLXINT held constant

	No Sleep Interruption	Sleep Interruption
N	138	24
$\log(\hat{P}/1-\hat{P})$	$-6.67 + 0.09\ L_{dn}$	$-4.80 + 0.07\ L_{dn}$
t	3.71 3.39	1.18 1.22
ρ^2	.14	.05
ρ_c^2	.07	.05
χ^2	26.35	1.79
χ_c^2	13.36	1.63
$\hat{E}N_{.5}$	70.52	67.5
σ	2.77	6.24

d = 0.44

disturbance levels under certain conditions of the intervening variables. The models for both of the groups which experienced activity interruption fail to give a significant fit to the data, while the coefficient of L_{dn} is non-significant in both cases. Again, this result may be due either to insufficient variation in noise levels (both of the groups which experienced activity interruption are exceedingly small); or to the presence of what Schultz calls "hypersensitives", these, the opposite of "imperturbables", being individuals who tend to report disturbance regardless of the noise level. Whatever the reason, it is impossible to draw definite conclusions about the effect of these two variables on disturbance from the data in Tables 5.5 and 5.6 alone.

Allowing both sensitivity and work-place noise rating to vary overcomes the problem of lack of fit (Tables 5.7 and 5.8) so that comparisons of the median effective noises can be made. However we again encounter interpretational problems because of the statistical relationship between the activity interruption variables on the one hand and the psychological variables *SENSE* and *RATWORK* on the other. Previously it has been argued that in this relationship the psychological variables correspond to cause and the activity interference variables to effect. Consistency therefore demands that the large inter-group variation in median effective noise in Table 5.7 be attributed in part to variations in sensitivity and rating of work-place noise. In a sense the problem is one of semantics: it is correct to say that large differences in the $EN_{.5}$ are *associated* with differences in relaxation interruption; but if it is desired to make a statement about the *cause* of those differences we need to be more careful.

Bearing in mind this caveat, note that the data for *RLXINT* (Table

Table 5.7

Differences in NHDATT (1-5/6-9) according to relaxation interruption: SLPINT held constant

	No Relaxation Interruption	Relaxation Interruption
N	555	101
$\log(\hat{P}/1-\hat{P})$	$-6.16 + 0.07\ L_{dn}$	$-4.62 + 0.07\ L_{dn}$
t	5.93 4.52	2.19 2.39
ρ^2	.36	.08
ρ_c^2	.04	.05
χ^2	278.63	10.73
χ_c^2	20.47	6.33
$\hat{E}N_{.5}$	85.84	62.35
σ	4.87	3.59

d = 3.88

Table 5.8

Differences in NHDATT (1-5/6-9) according to sleep interruption: RLXINT held constant

	Sleep Not Interrupted	Sleep Interrupted
N	101	86
$\log(\hat{P}/1-\hat{P})$	$-4.62 + 0.07\ L_{dn}$	$-4.53 + 0.08\ L_{dn}$
t	2.19 2.39	2.15 2.45
ρ^2	.08	.13
ρ_c^2	.05	.06
χ^2	10.73	15.98
χ_c^2	6.33	6.69
$\hat{E}N_{.5}$	62.35	59.51
σ	3.59	4.54

d = 0.49

5.7) are consistent with the hypothesized direction of explanation. Since relaxation interruption, itself a measure of noise impact, is more closely linked to the response variable than either *SENSE* or *RATWORK* its relationship to response is mitigated by fewer intervening factors, and, as we would expect, its effect on the median effective noise is even more marked than that of the psychological variables.

The same is not true however of *SLPINT* (Table 5.8), but the explanation of this finding lies in the relatively high correlation of the two activity interruption variables (phi-coefficient of .4). If *RLXINT* is *not* controlled for in looking at *SLPINT* the d-statistic rises to 4.17, with individuals who suffer interruption of sleep being prone to disturbance at considerably lower noise levels than those who don't. This interrelationship between sleep interruption and relaxation interruption is further evidence to support the hypothesis that both types of activity interruption are influenced by underlying psychological variables such as sensitivity. The finding is also consistent with previous evidence for the intercorrelation of activity interference variables (see Section 4.1).

5.2 Differences in Median Effective Noise for Severe Disturbance by Overall Neighbourhood Noise

On the basis of the results in Section 4.2, three variables will be examined here for their effect on inter-group variations in median effective noise levels: sensitivity, rating of work-place noise, and conversation interruption.

As was the case with the analysis for *NHDATT (1-5/6-9)*, calibrations of the model

$$P = (1 + e^{-\alpha - \beta L_{dn}})^{-1}$$

for particular cells of the 3×3 matrix of individual attributes resulted in non-significance of the goodness of fit statistics and the coefficients of L_{dn} (Tables 5.9-5.11). Such models continue to suggest that at certain values of the vector of intervening variables, response, whilst partly determined by the levels of these attributes, is independent of the noise level.

Analysis for *SENSE* and *RATWORK* proceeded by allowing *CONINT* to vary between groups. Still, the noise level completely fails to explain response for the less sensitive individuals and those with relatively quiet work-places (Tables 5.12, 5.13) resulting in very large standard errors in the $EN_{.5}$ estimates. In order to obtain low variance estimates and significant overall fits to the data it was necessary, when testing a particular intervening variable, to relax control of all other variables. The results for both sensitivity and work-place noise rating (Tables 5.14, 5.15) reveal relatively small variations in $EN_{.5}$ levels along the dimensions of these two variables. In part this is due to the persistence of large variation in the $EN_{.5}$ estimates for the two "base" groups - i.e. the less sensitive individuals and the individuals with relatively quiet work-places. This in turn appears to result from the presence of "imperturbables" in each group as is suggested by the fact that 97.6% of the less sensitive group and 97.5% of those who reported quieter than average work-places were *not* severely disturbed. In such situations the concept of a median effective noise has little meaning and is of doubtful validity. Certainly

Table 5.9

Differences in NHDATT (1-7/8-9) according to

sensitivity: RATWORK and CONINT held constant

	Less Sensitive Group	More Sensitive Group
N	84	113
$\log(\hat{P}/1-\hat{P})$	$-7.37 + 0.05\ L_{dn}$	$-9.70 + 0.12\ L_{dn}$
t	1.13 0.57	4.16 3.69
ρ^2	.84	.36
ρ_c^2	.02	.14
χ^2	97.88	55.93
χ_c^2	0.33	16.14
$\hat{E}N_{.5}$	134.21	78.52
σ	115.91	3.23

d = 0.48

Table 5.10

Differences in NHDATT (1-7/8-9) according to rating of work-place noise: SENSE and CONINT held constant

	Less Noisy Work-place	More Noisy Work-place
N	261	84
$\log(\hat{P}/1-\hat{P})$	$-10.27 + 0.09\ L_{dn}$	$-7.37 + 0.05\ L_{dn}$
t	1.97 1.16	1.13 0.57
ρ^2	.91	.84
ρ_c^2	.04	.02
χ^2	330.32	97.88
χ_c^2	1.26	0.33
$\hat{EN}_{.5}$	115.42	134.21
σ	41.58	115.91

$d = -0.15$

Table 5.11

Differences in NHDATT (1-7/8-9) according to conversation interruption: SENSE and RATWORK held constant

	No Conversation Interruption	Conversation Interruption
N	113	31
$\log(\hat{P}/1-\hat{P})$	$-9.70 + 0.12\ L_{dn}$	$-3.09 + 0.07\ L_{dn}$
t	4.16 3.69	0.66 1.04
ρ^2	.36	.47
ρ_c^2	.14	.05
χ^2	55.93	20.29
χ_c^2	16.14	1.15
$\hat{E}N_{.5}$	78.52	44.59
σ	3.23	25.99

d = 1.30

Table 5.12

Differences in NHDATT (1-7/8-9) according to

sensitivity: RATWORK held constant

	Less Sensitive Group	More Sensitive Group
N	94	114
$\log(\hat{P}/1-\hat{P})$	$-8.36 + 0.08\ L_{dn}$	$-9.68 + 0.13\ L_{dn}$
t	1.76 1.15	5.16 4.75
ρ^2	.76	.29
ρ_c^2	.04	.17
χ^2	98.59	57.60
χ_c^2	1.36	28.13
$\hat{E}N_{.5}$	107.85	77.01
σ	33.86	2.15

d = 0.91

Table 5.13

Differences in NHDATT (1-7/8-9) according to rating of work-place noise: SENSE held constant

	Quieter Work-place	Noisier Work-place
N	284	114
$\log(\hat{P}/1-\hat{P})$	$-8.13 + 0.07\ L_{dn}$	$-9.68 + 0.13\ L_{dn}$
t	2.73 1.63	5.16 4.75
ρ^2	.80	.29
ρ_c^2	.32	.17
x^2	316.41	57.60
x_c^2	2.54	28.13
$\hat{E}N_{.5}$	112.94	77.01
σ	28.35	2.15

d = 1.26

Table 5.14

Difference in NHDATT (1-7/8-9) according to sensitivity

	Less Sensitive Group	More Sensitive Group
N	373	428
$\log(\hat{P}/1-\hat{P})$	$-12.40 + 0.13\ L_{dn}$	$-11.11 + 0.13\ L_{dn}$
t	3.92 2.91	7.63 6.50
ρ^2	.85	.57
ρ_c^2	.10	.16
χ^2	440.93	337.39
χ_c^2	8.66	48.61
$\hat{E}N_{.5}$	96.44	83.43
σ	9.10	2.32

d = 1.38

Table 5.15

Differences in NHDATT (1-7/8-9)

according to rating of work-place noise

	Quieter Work-place	Noisier Work-place
N	563	238
$\log(\hat{P}/1-\hat{P})$	$-10.46 + 0.10\ L_{dn}$	$-9.97 + 0.12\ L_{dn}$
t	4.36 2.97	6.03 5.31
ρ^2	.84	.41
ρ_c^2	.06	.15
χ^2	657.84	135.57
χ_c^2	8.44	33.49
$\hat{E}N_{.5}$	102.12	81.43
σ	11.45	2.43

d = 1.77

the estimated $EN_{.5}$ represents extrapolation well beyond the range of the data: hence the rather unlikely value of 102 dBA obtained for the quieter work-place group (Table 5.15) and the large standard error of 11.45 dBA associated with it.

The results for conversation interruption (Table 5.16) are somewhat better, in respect of the variance of the parameter estimates. Surprisingly, the data indicate little difference in median effective noise levels caused by interruption of conversation. The apparent contradiction of earlier findings by this result can be explained by focussing on the parameters of the response equation. As pointed out earlier, (Chapter 2, Section 2.3), entering non-acoustical variables directly into the response equations allows the effect of these variables to emerge only through the constant term: the coefficient of noise is constrained to remain the same regardless of the value of $\underline{\underline{Z}}$, or in this case of conversation interruption. Now the constant term *does* in fact vary widely according to the value of *CONINT* (Table 5.16); hence the significance of the variable in earlier model estimations (Table 4.8). However, the present analysis reveals that the coefficient of noise also varies between groups, and in fact the difference in β for the two groups has an effect on the response probabilities opposite that of α. Whereas at noise levels approaching 0 the probability of response is greater for individuals, in the interrupted group, the tolerance distribution for the individuals reporting no interruption clusters more tightly about its mean so that as the noise level increases the difference between the cumulative response probabilities decreases quite rapidly. Thus the d-statistic for the $EN_{.2}$ group for example is 2.65, almost 3 times that associated with the median effective

Table 5.16

Differences in NHDATT (1-7/8-9) according to

conversation interruption

	No Conversation Interruption	Conversation Interruption
N	714	87
$\log(\hat{P}/1-\hat{P})$	$-11.20 + 0.12\ L_{dn}$	$-6.69 + 0.08\ L_{dn}$
t	6.69 5.12	2.79 2.46
ρ^2	.76	.19
ρ_c^2	.10	.06
χ^2	749.09	22.89
χ_c^2	26.63	6.64
$\hat{E}N_{.5}$	91.33	84.48
σ	4.52	5.30

d = 0.98

noise. This result helps to demonstrate that the assessed difference between any two groups with respect to their tolerance distributions (response curves) depends upon the response level in which we are interested.

5.3 Differences in Median Effective Noise for Any Disturbance by Main Road Traffic

Two non-acoustical variables have been identified as significantly related to *MNATT (0–5/6–9)* - sensitivity and level of education.

Looking first at the results for sensitivity (Table 5.17) the analysis reveals that there is a large difference in the median effective noise between the more and less sensitive sections of the population. As we might expect, the average effective noise, or tolerance, among the more sensitive individuals is 9.08 dBA higher than that for the less sensitive individuals. Moreover the β parameters of the tolerance distributions indicate that the probability of response increases more quickly with noise in the sensitive group. This result is interesting in so far as it is common to find reference to β as a "sensitivity" parameter; that is a parameter which indicates how "sensitive" or responsive the individual is to changes in the value of the independent variable, be it noise level, the cost of travel, or the price of a good (Wilson, 1974, 143). The present finding gives real and precise meaning to this concept. Moreover, the difference in β between the groups emphasizes again the qualitative difference between this and the earlier multivariate logit analyses, and suggests that the assumption of a constant noise parameter on which the multivariate model is based may not be tenable.

Dividing the sample on the basis of level of education, and con-

Table 5.17

Differences in MNATT (0-5/6-9) according to sensitivity: LEVED held constant

	Less Sensitive Group	More Sensitive Group
N	193	267
$\log(\hat{P}/1-\hat{P})$	$-11.38 + 0.14\ L_{dn}$	$-11.11 + 0.16\ L_{dn}$
t	4.84 4.17	6.94 6.55
ρ^2	.52	.26
ρ_c^2	.14	.18
χ^2	139.54	96.05
χ_c^2	20.79	59.99
$\hat{E}N_{.5}$	80.14	71.06
σ	3.17	1.15

d = 2.69

trolling for sensitivity, yields large inter-group differences in $EN_{.5}$ (Table 5.18). The data indicate that a response level of 50% disturbed is reached at a lower noise level among the more highly educated individuals than it is among individuals with a lower level of educational attainment. The relatively high d-statistic confirms the significance of this variable in earlier analyses.

5.4 Differences in Median Effective Noise for Severe Disturbance by Main Road Traffic Noise

Only one intervening variable, sensitivity, has been found to be significantly related to the incidence of disturbance by main road traffic noise. Comparison of the mean values ($EN_{.5}$'s) of the fitted tolerance distributions for more and less sensitive individuals reveals however a relatively small difference (in standard deviation units) between these two groups (Table 5.19). Yet again this result is indirectly due to the fact that only 1% of the individuals in the less sensitive group reported severe disturbance.

The poor explanatory power of L_{dn} in the model for the less sensitive group, as compared to its relatively good performance in the models for the median cut-off (*MNATT (0-5/6-9)*), describes a trend similar to that found with analysis of *NHDATT*. Whereas noise level usually explains *any* disturbance fairly well within specific groups of the sample (though it may be necessary to aggregate over certain variables, as in Section 5.1), greater difficulties have been experienced in analyzing *severe* disturbance. Schultz claims that severe disturbance is better explained by noise alone than is median disturbance, and the earlier

Table 5.18

Differences in MNATT (0-5/6-9) according to level of education: SENSE held constant

	Only Some High School	At Least High School Graduation
N	176	267
$\log(\hat{P}/1-\hat{P})$	$-8.23 + 0.10\ L_{dn}$	$-11.11 + 0.16\ L_{dn}$
t	5.04 4.37	6.94 6.55
ρ^2	.34	.26
ρ_c^2	.17	.18
χ^2	81.82	96.05
χ_c^2	21.46	59.99
$\hat{E}N_{.5}$	79.39	71.06
σ	3.15	1.15

d = 2.49

Table 5.19

Differences in MNATT (0-7/8-9) according to sensitivity

	Less Sensitive Group	More Sensitive Group
N	387	445
$\log(\hat{P}/1-\hat{P})$	$-15.15 + 0.15\ L_{dn}$	$-11.93 + 0.14\ L_{dn}$
t	3.16 2.37	8.00 6.92
ρ^2	.93	.59
ρ_c^2	.13	.18
χ^2	497.84	361.17
χ_c^2	5.88	57.03
$\hat{E}N_{.5}$	98.27	82.77
σ	11.07	2.03

d = 1.38

analysis supported this assertion (c.f. Chapter 4, Section 4.5). However, the present finding serves as a warning that the matter is a good deal more complex than Schultz's statement suggests, since within particular groups of individuals noise performs very poorly in explaining response - worse, in fact, than non-acoustical variables. For example, the present result implies that if we know an individual is relatively less sensitive to noise, we can predict whether or not he will be severely disturbed almost with certainty, and without any knowledge of the relevant noise level.

5.5 Summary and Discussion

This chapter has extended the multivariate analysis of response by comparing estimated tolerance distributions across different groups of the sample defined according to their levels of the intervening variables shown in Chapter 4 to be significantly related to response. The comparisons have been based upon the $EN_{.5}$, which for a given sample or group corresponds to the mean and median of the tolerance distribution. The use of this measure to test for differences in response takes into account simultaneously variation in both parameters of the tolerance distribution (response curve). The analysis is therefore qualitatively different from that undertaken in Chapter 4, with the result that some of the results appear to contradict earlier findings. For example, conversation interruption was found in the first part of the analysis to be significantly related to severe disturbance by neighbourhood noise levels; but the present analysis suggests little difference in median effective noise due to this variable.

Other important results of the analysis include the finding that sleep interruption causes little difference in the $EN_{.5}$, most of the correlation of this variable with response deriving from its intercorrelation with relaxation interruption. Sensitivity emerges as the most important determinant of the median effective noise level, not only because of its direct effect on the $EN_{.5}$, but also because of its indirect effect through other variables such as the activity interruption measures. It is argued that sensitivity is, in a sense, the *causa causans* in the relationship between exposure and disturbance. The only socio-economic variable previously found to be significantly related to response, level of education, is also associated with relatively large differences in the median effective noise.

More important than these results for particular intervening variables however is the finding that in many instances response is actually independent of the noise level. The analysis of *NHDATT (1-5/6-9)*, for example, suggests that the effect of the four intervening variables on response so outweighs that of L_{dn} that when all four are held constant simultaneously most of the variation in response is removed; and because noise fails to explain the remaining variation its coefficient does not achieve significance at acceptable confidence levels. This finding supports the central thesis underlying the present work; namely, that non-acoustical variables play at least as great a part as L_{dn} in determining individual response to noise.

The $EN_{.5}$, though a useful summary measure of response, probably represents too high a level of disturbance (or severe disturbance) to be acceptable as a basis for the setting of noise standards. There is, however, no commonly accepted "allowable percentage disturbed", and though Schultz acknowledges that zero disturbance is impossible most authors who have been concerned with design noise levels have avoided committing themselves on this matter. The study by Schultz does however identify a noise level of 65 dBA L_{dn} as an optimal standard if the number of people highly annoyed is to be minimized subject to certain feasibility constraints. Given the noise-response curve on which Schultz bases his study, this noise level corresponds to approximately 15% highly annoyed.

If we accept 15% highly annoyed (or disturbed), as something close to the level of disturbance which might be considered as acceptable for a noise abatement program incorporating some feasibility considerations, we can calculate the noise standards ($EN_{.15}$) which would be needed to achieve this level of response according to various studies. For the present study, $EN_{.15}$'s were calculated for the two variables which measure severe disturbance - *NHDATT (1-7/8-9)* and *MNATT (0-7/8-9)*. Similar response levels were calculated from the response equation derived from an earlier study which used the same data set as the present analysis (Hall, Taylor and Birnie, 1977), while data from a French motorway study described by Schultz were also included in the analysis (Lamure and Bacelon, 1976). All of these surveys were conducted at the aggregate level; they all attempted to identify the best fit curve for plots of percentage highly disturbed (or highly annoyed) against noise levels.

The data for these four studies (Table 5.20) indicate strong agreement between the study conducted by Hall *et al.*, and the present analysis, particularly with respect to the $EN_{.15}$ for *NHDATT*. It would appear therefore that the level of aggregation at which the analysis is conducted has little bearing on the results obtained - in terms of the $EN_{.15}$ at any rate.

Table 5.20

Comparison of noise levels (L_{dn}) required to produce 15% highly annoyed/disturbed in various studies

Present Study	Hall *et al.*	Schultz	Lamure and Bacelon
74.1 (NHDATT)	75.47 (NHDATT)	65	56.3
74.9 (MNATT)	68.96 (MNATT)		

On the other hand there is a considerable difference between the results of the present study and those of Schultz and Lamure. For example the level of 74.9 dBA which represents the $EN_{.15}$ for *MNATT (0-7/8-9)* would correspond to 36.59% highly annoyed in Schultz's response equation and to 75.9% highly annoyed in the equation reported by Schultz for the French study. Given that the discrepancy between these figures is unlikely to be related to the aggregation issue, the most obvious single explanation is to attribute it to differences in the samples upon which the studies are based. The results are consistent with this reasoning in that, of the two studies conducted on "other" data sets, Schultz's shows least divergence from the present study. This is to be expected since the noise-response curve derived by Schultz represents an averaging or synthesis of the results of 11 other studies, (including the Lamure and Bacelon study) so that differences in sample composition between it and the present study could be expected to be less.

If it is correct to impute the differences in these "potential noise standards" to differences in the attributes of the observations (individuals) on which the studies are based, the conclusion must be that design noise levels, to be successful, cannot be considered independently of the sample from which they were derived nor of the population to which they are to be applied. The results of the present study are to some extent their own justification therefore, in so far as they indicate the need for an approach which attempts to identify explicitly the effects of sample characteristics.

CHAPTER 6

SUMMARY AND CONCLUSIONS

The objective of this thesis has been to develop a model of individual response to community noise in order to analyze the effect of non-acoustical variables on response. The study was motivated by a desire to assess the extent to which it is possible to formulate effective noise control policy in terms of noise levels alone without consideration of the characteristics of the exposed population. The purpose of this chapter is to summarize the substance of each of the preceeding chapters; to evaluate the results of the analysis in terms of the research objectives; and to suggest avenues for further research stemming from the present study.

6.1 Summary

Chapter 1 contained the problem statement and outlined the research objectives. It was pointed out that existing noise control regulations have been drawn up on the assumption that response to noise is largely, if not completely, determined by noise levels alone. Previous studies of response were shown to have ignored for the most part questions of the effect of non-acoustical, "intervening" variables in the relationship between noise and response. The failure of these studies to reach agreement about the nature of this relationship was cited as evidence to support the hypothesis that attributes of respondents have in fact a significant effect on response. The lack of precise information on the effect of respondent characteristics was related to the failure to identify an appropriate technique for modelling individual response.

Chapter 2 invoked quantal response theory in order to develop a formal model of response to noise. The key element of this theory was shown to be the concept of a threshold stimulus level corresponding, in the case of noise, to the maximum noise level an individual can tolerate - his tolerance - before a quantum change in his response occurs. Different assumptions about the distribution of individual tolerances in the population were related to particular forms of response model and reasons for the selection of the logit model outlined. Two methods of dealing with response in heterogeneous populations were described and the difference between these was shown to be related to their assumptions about the relationship between the tolerance distributions of different kinds of individuals. Extensions of the binary response model were considered, but dismissed as being more cumbersome than useful.

Chapter 3 was concerned with operationalizing the model of response. The abstract components of the model were translated into specific empirically-measured variables for purposes of analysis. Following from the theoretical discussion in Chapter 2, two methods of examining the effects of non-acoustical variables on response were discussed, the advantages and disadvantages of each evaluated, and a decision made to use both in the analysis. The chapter concluded with a description of the statistics used to guide the analysis.

The purpose of Chapter 4 was to derive the best predictive equations for each of the 4 response variables in turn. Among the more important results were the findings that: (i) noise level alone gives a very poor explanation of response; (ii) the incidence of severe disturbance is easier to predict than that of any disturbance at all; and (iii) attitude towards main road traffic noise is easier to predict than attitudes towards overall neighbourhood noise. Of the intervening variables sensitivity was found to give the best explanation of response while the socio-economic and demographic variables, with the exception of education, performed poorly in this respect. In general, overall explanation of response was low for all four of the dependent variables.

Chapter 5 described the results of calibrating response models for groups of individuals which were internally homogeneous with respect to the more important intervening variables. The most important result of this analysis was the finding that for certain groups response is independent of the noise level. Quite large inter-group differences in the average tolerance level were found to be associated with differences in the individual's sensitivity and his rating of the noise level at his place of work. Surprisingly, level of education was also important in this respect. Finally, comparison of the results of the present study with those of other researchers revealed large differences in the noise levels required to cause 15% (chosen because of its possible significance for design noise levels) of the respective samples to be highly annoyed/disturbed. These differences were tentatively imputed to differences in characteristics of the individuals making up the various samples.

6.2 Conclusions

Corresponding to the twin objectives of the thesis the major conclusions can be separated into those which refer primarily to the development of the model of response; and those which have to do with the substantive results of the analysis.

With regard first to the model used to analyze response it is felt that, despite poor empirical results (in terms of goodness of fit), the modelling approach adopted here has much to recommend it. Its major strength lies in the fact that it is derived directly from a theory of response, so that even specific elements of the mathematical form used (in the present case the logistic function) can be related back to fundamental theoretical concepts. In addition, the interpretation and explanation of results for specific variables can be given extra depth by reference to the tolerance concept. Thus we do not just observe that individuals with, say, noisier work-places, tend to be more disturbed than others by neighbourhood noise; rather, the use of the tolerance idea involves implicitly an explanation as to why this might be so.

It must be recognized however that if the strengths of the present modelling approach lie in its theoretical and conceptual basis, it is on the same grounds that it is most open to criticism. Thus questions about the suitability or "goodness" of the model - other than questions about the specific functional form - reduce to questions about the merit of the tolerance concept. Unfortunately discussions of this point are apt to become dogmatic and unfruitful. However if we keep our terms of reference empirical rather than theoretical it seems reasonable to suggest that the tolerance idea is likely to work better for certain variables than for others.

Undoubtedly the concept makes more sense in conjunction with a variable like sensitivity (with which it is almost synonymous), than it does in conjunction with a variable like income: that is, to go back to the point about "implied explanations" made earlier, the finding that highly sensitive individuals are, *ceteris paribus*, more likely to be disturbed than non-sensitive individuals is clearly more readily explained by the tolerance concept than is a (hypothetical) finding that higher income tends also to predispose individuals to disturbance. One wonders therefore whether the relatively good performance of sensitivity, and the poor performance of income (and similar socio-economic variables) in the models calibrated here, might not be due to differences in the suitability of the tolerance idea. Further analysis might investigate this idea further.

Turning to consider the effect of non-acoustical variables on response, three major conclusions can be drawn from the analyses. First, despite the wide array of intervening variables analyzed, overall explanation of response was consistently low, forcing the conclusion that the determinants of individual response to noise remain largely unknown. Secondly, noise level alone is in general a very poor predictor of response; and thirdly, whereas psychological variables such as sensitivity or rating of work-place noise contribute significantly to explanations of response, socio-economic and demographic variables perform poorly. The single exception to the latter statement is level of education which does exhibit some relationship to disturbance. Though the generally poor levels of explanation may possibly be due to the unsuitability of the model of response used, as discussed above, it is more likely that they result from gaps in the data on individual characteristics. It is too early to resign the attempt to explain individual response and future research might usefully investigate characteristics of individuals other than those treated here. On the basis of the present results, investigation of other psychological variables would appear to be potentially the most fruitful approach, though the stability of the observed relationship between education and response might also be tested. The poor explanation of response afforded by the noise metric is probably due in part to shortcomings in the noise measurements used. In this respect it is important to note that the data on which the present analyses are based comprise noise readings for entire residential sites and not individual respondents. Although one of the criteria governing site selection was that the noise level should be more or less uniform within the site, it is clear that these aggregate data are less than optimal for use in a disaggregate model of response.

As regards the question of whether effective noise control policy can be formulated by considering noise levels alone, an example has already been provided (see Chapter 5, Section 5.5) to illustrate the wide divergence that exists between studies with respect to their descriptions of the noise-response relationship. Unfortunately the disagreements among these studies could only be postulated to be due to differences in sample composition, and there are insufficient data to test this hypothesis. Confining ourselves to what data there are available we can only give a conditional answer to the question. In effect the answer depends upon whether or not the distributions of variables such as sensitivity and rating of work-place noise are spatially or temporally autocorrelated. That is, if, for example, there is a strong tendency for highly sensitive individuals to be found at quiet residential sites, and less sensitive individuals at more noisy sites then the answer to the question is No. We need only look at the differences in median effective noise due to sensitivity to realize this. In terms of the example given, either design noise levels would have to be set low enough to be acceptable to the highly sensitive individuals; or else different noise standards would need to be set for the two sites, in recognition of the differences between the two groups of individuals. If it had transpired that, say, income was significantly related to response, then in view of the tendency for the spatial distribution of income to conform to definite patterns we could have rejected categorically the idea of planning in terms of the noise alone. But since it would appear that sensitivity, and not income, is the relevant variable, since we know little about the spatial distribution of sensitivity, and since in any case the larger part of the variation in response remains unaccounted for by any of the variables examined here, the issue of planning in terms of noise alone cannot be satisfactorily resolved. The direction of further research should however be obvious.

APPENDIX 1

THE QUESTIONNAIRE

This appendix contains a copy of the questionnaire used during the first summer of the survey (1975). This version is presented instead of the slightly different 1976 questionnaire due to the fact that most of the analyses were based upon the 1975 data set (see Chapter 4, Section 4.1).

FINAL QUESTIONNAIRE

SITE NUMBER ________ (1-4)

RESPONDENT NUMBER ________ (5-7)

[*INTRODUCTION: Hello, I'm from the Geography Department at McMaster and I'm interviewing people to find out what they think about this area. Could you spare me about 10 minutes? Thanks very much.*]

1. a. What are the important things you like about living in this neighbourhood?

   b. Which of these is the most important to you?

2. a. What are the important things you don't like about living in this neighbourhood?

   b. Which of these is the most important to you?

Item		Like V (1)	Like E (2)	Important	Don't Like V (1)	Don't Like E (2)	Important
(01)	Schools	8			28		
(02)	Shopping	9			29		
(03)	Open Space	10			30		
(04)	Recreational Facs.	11			31		
(05)	Bus Service	12			32		
(06)	Proximity to Work	13			33		
(07)	Noise	14			34		
(08)	Quietness	15			35		
(09)	Air Quality	16			36		
(10)	Landscaping	17			37		
(11)	Cost of Housing	18			38		
(12)	Quality of Housing	19			39		
(13)	Neighbours	20			40		
(14)	Safety for Children	21.			41		
(15)	Crime	22			42		
(16)	Other (Specify)	23		24-27*	43		44-47*

*[*When coding put ID of most important in 24-25 (44-45), and use additional two cols. if a second source was mentioned.*]

[*For 1 and 2 check all non-volunteered items using card ie. Here are some other things that have been mentioned, are any of them important to you?*]

[*Transition: You have mentioned noise; I'd like to ask you a little more about that. (or) One of the items we're particularly interested in is noise and I'd like to ask you about that.*]

[*People tend to vary in their sensitivity to noise....*]

3. In general, how sensitive are you to noise?

(1) Not at all ☐

(2) A little ☐

(3) Moderately ☐ (48)____

(4) Considerably ☐

(5) Extremely ☐

[*Number of noise sources mentioned.*] (49)____

[*When coding skip to question 5.*]

4. a. What noises are clearly audible to you in this neighbourhood?

   b. How would you rate each of the noises you've mentioned?
   [Hand respondent the card listing intensity scale.]

   c. Here is a list of common noises. (You have already mentioned some.)
   Do you ever notice any of these (any of the others)?

   d. *[Repeat* b *for elicited noises.]*

   e. *[For each noise with an intensity score between 6 and 9:]*
   What aspects of these noises do you find disturbing
   (loudness, pitch, intermittent nature, suddenness, variation)?

   [Coding: Each source mentioned will appear on a separate card. Duplicate the identification code in cols. (1-7). Noise source code goes in cols. (8-9). Proceed with data from questions 4, 8, 9, 10, 11, and 13 for that noise source. Repeat as needed for additional noise sources.]

Source	Vol.	Elic.	Intensity	Aspects (No=1			Yes=2)	
	1	2	1-9	L	P	I	S	V
(8-9)	(10)		(11)	(12)	(13)	(14)	(15)	(16)
(01) Children	___	___	___	___	___	___	___	___
(02) Other People	___	___	___	___	___	___	___	___
(03) Handyman Tools	___	___	___	___	___	___	___	___
(04) Air Conditioner	___	___	___	___	___	___	___	___
(05) Domestic Pets	___	___	___	___	___	___	___	___
(06) Garden Machinery	___	___	___	___	___	___	___	___
(07) TV/Radio/Records	___	___	___	___	___	___	___	___
(08) Musical Instruments	___	___	___	___	___	___	___	___
(09) Local Traffic Noise	___	___	___	___	___	___	___	___
(10) Freeway Traffic Noise	___	___	___	___	___	___	___	___
(11) Motorcycles	___	___	___	___	___	___	___	___
(12) Trucks	___	___	___	___	___	___	___	___
(13) Snowmobiles	___	___	___	___	___	___	___	___
(14) Mini-bikes	___	___	___	___	___	___	___	___
(15) Trains	___	___	___	___	___	___	___	___
(16) Aircraft	___	___	___	___	___	___	___	___
(17) Industrial Noise	___	___	___	___	___	___	___	___
(18) Construction Noise	___	___	___	___	___	___	___	___
(19) Institutional Noise	___	___	___	___	___	___	___	___
(20) Other (Specify)	___	___	___	___	___	___	___	___
(8-9)	(10)		(11)	(12)	(13)	(14)	(15)	(16)

5. a. Did any of the noises listed on the card which don't disturb you now, ever disturb you or threaten to disturb you in the past?

______________________ (50-51) _____

______________________ (52-53) _____

______________________ (54-55) _____

5. b. Why are these noises no longer (potentially) disturbing?

	SOURCES		
	______	______	______
EXTERNAL AGENCIES			
(11) Newspaper	☐	☐	☐
(12) Noise source	☐	☐	☐
(13) Police	☐	☐	☐
(14) Politicians	☐	☐	☐
(15) Other gov't. officials	☐	☐	☐
(16) Protest group	☐	☐	☐
(17) Other (specify)	☐	☐	☐
PERSONAL ACTIONS			
(21) Got used to noise	☐	☐	☐
(22) Installed extra insulation	☐	☐	☐
(23) Installed double glazing	☐	☐	☐
(24) Planted trees	☐	☐	☐
(25) Moved to a quieter neighborhood	☐	☐	☐
(26) Other (specify)	☐	☐	☐
	______	______	______
	71-72	73-74	75-76

[*When coding use cols. (71-76) on card 2.*]

6. How would rate the overall noise in this neighbourhood?

(1) Extremely Agreeable ☐

(2) Considerably Agreeable ☐

(3) Moderately Agreeable ☐

(4) Slightly Agreeable ☐

(5) Neutral ☐ (56)____

(6) Slightly Disturbing ☐

(7) Moderately Disturbing ☐

(8) Considerably Disturbing ☐

(9) Extremely Disturbing ☐

7. [*Ask only if traffic, trucks or trains have been mentioned as disturbing.*]

You have mentioned that__________ is/are disturbing to you. If it were somehow possible to remove the noise, do you think the value of this (house or apartment) would increase? By how much?

(1) No ☐

(2) Yes ☐ (57)____

(3) Don't know ☐

House value	$______________	(58-60)
	%______________	(61-62)
Apartment value	$______________	(63-64)
	%______________	(65-66)

[*When coding, skip to question 12.*]

[*Accept whichever type of response is offered, dollars or percentage, do not try to get both.*]

8. a. What time of the year (summer or winter) are you disturbed most by these noises?

b. Are you disturbed most on weekends (Sat. or Sun.) or during the week?

c. What time of day does each noise disturb you most?

Source	Annual		Weekly				
	S	W	W	W/E	W+S	S	SN
	(1)	(2)	(1)	(2)	(3)	(4)	(5)

____________ ☐ (17) ☐ (18)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 (W) (19-26)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 (S) (27-34)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 (SN) (35-42)

____________ ☐ (17) ☐ (18)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 (W) (19-26)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 (S) (27-34)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 (SN) (35-42)

____________ ☐ (17) ☐ (18)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 (W) (19-26)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 (S) (27-34)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 (SN) (35-42)

____________ ☐ (17) ☐ (18)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 (W) (19-26)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 (S) (27-34)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 (SN) (35-42)

9. a. Where are you when each of these noises disturbs you?

b. What are you doing when these noises disturb you?

Source	Location				Activity							
	BY	FY	IN	EW	SL	RL	RD	CV	WK	TV	TC	E
______	__	__	__	__	__	__	__	__	__	__	__	__
______	__	__	__	__	__	__	__	__	__	__	__	__
______	__	__	__	__	__	__	__	__	__	__	__	__
______	__	__	__	__	__	__	__	__	__	__	__	__
	43	44	45	46	47	48	49	50	51	52	53	54

[*Blank = 1, check = 2*]

10. What effects on the health of you and your family do you think continued exposure to these noise sources could have?

Source	None	Nervousness	Hearing loss	Irritability	Headaches	Inter. Sleep	Kept Awake
______	__	__	__	__	__	__	__
______	__	__	__	__	__	__	__
______	__	__	__	__	__	__	__
______	__	__	__	__	__	__	__
	55	56	57	58	59	60	61

[*Blank = 1, check = 2*]

11. When you are disturbed by ____________ do you:

	Sources			
	______	______	______	______
Do nothing	☐	☐	☐	☐
Close your window	☐	☐	☐	☐
Use air conditioning	☐	☐	☐	☐
Stay indoors	☐	☐	☐	☐
Turn on/up T.V./radio/records	☐	☐	☐	☐
Wear earplugs	☐	☐	☐	☐
Contact noise source	☐	☐	☐	☐
Wait for noise to stop	☐	☐	☐	☐

12. Who would you contact if you had a specific complaint about noise?

Don't know	☐	(67)
Noise source	☐	(68)
Police	☐	(69)
MOE	☐	(70)
Politician	☐	(71)
Other govt. official	☐	(72)
Protest group	☐	(73)
Other (specify)	☐	(74)

[Blank = 1, check = 2]

[When coding, skip to question 14.]

13. Have you ever taken any of these actions in response to these noise sources?

	Sources				
	______	______	______	______	
Done nothing	☐	☐	☐	☐	70
Written to newspaper	☐	☐	☐	☐	71
Contacted noise source	☐	☐	☐	☐	72
Contacted police	☐	☐	☐	☐	73
Contacted politician	☐	☐	☐	☐	74
Contacted other govt. official	☐	☐	☐	☐	75
Signed petition	☐	☐	☐	☐	76
Attended meeting	☐	☐	☐	☐	77
Joined protest group	☐	☐	☐	☐	78
Organized protest group	☐	☐	☐	☐	79
Other (specify)	☐	☐	☐	☐	80

[Blank = 1, check = 2]
[When coding, if there are more sources, return to question 4.]

14. Have you considered any long-term actions to avoid unwanted noise? If so, which of the following have you considered? Have you ever taken any of these actions.

	Cons.	Taken
None	☐ (75)	☐ (76)
Install extra insulation	☐ (77)	☐ (78)
Install air conditioning	☐ (79)*	☐ (8)
Install double glazing	☐ (9)	☐ (10)
Plant trees	☐ (11)	☐ (12)
Move to a quieter neighbourhood	☐ (13)	☐ (14)
Other (specify)	☐ (15)	☐ (16)

[*Blank = 1, check = 2*]

*[*When coding, col. (80) = 1; start a new card by duplicating the identification code in cols. (1-7); then punch cols. (8-16) as above, and proceed to question 15.*]

15. How would you rate this neighbourhood for noise compared with other residential parts of the Hamilton area?

(1) Very quiet ☐

(2) Fairly quiet ☐

(3) Average ☐ (17)____

(4) Fairly noisy ☐

(5) Very noisy ☐

16. Would you be willing to pay extra taxes in order to help control noise sources in your neighbourhood?

(1) No ☐ (18)____

(2) Yes ☐

(If yes) Approximately how much? $__________ (19-21)____

17. a. Do you think existing municipal by-laws are effective in controlling noise?

(1) No ☐

(2) Yes ☐ (22)____

(3) Don't know ☐

b. Do you think existing provincial/federal legislation is effective in controlling noise?

(1) No ☐

(2) Yes ☐ (23)____

(3) Don't know ☐

18. Do you have any suggestions for helping to reduce noise?

PERSONAL DATA

1. Sex:

   (1) Male ☐ (24)____

   (2) Female ☐

2. Please write down your age on your last birthday ________ years. (25-26)

3. What level of education have you completed?

   (1) Some public school ☐

   (2) Public school graduation ☐

   (3) Some high school ☐

   (4) High school graduation ☐ (27)____

   (5) Some university or college ☐

   (6) University or college graduation ☐

   (7) Post-graduate work ☐

4. What is your main occupation? ______________________ (28)____

5. How would you rate your place of work for noise?

   (1) Very quiet ☐

   (2) Fairly quiet ☐

   (3) Average ☐ (29)____

   (4) Fairly noisy ☐

   (5) Very noisy ☐

6. Please indicate which range most closely describes the before taxes income supporting this household in the past year.

(1) Less than $5,000 ☐ (5) $20,000 - $25,000 ☐

(2) $5,000 - $10,000 ☐ (6) $25,000 - $30,000 ☐ (30)____

(3) $10,000 - $15,000 ☐ (7) More than $30,000 ☐

(4) $15,000 - $20,000 ☐

7. How many hours do you normally spend at home each day? (out of 24)

______________________ on weekdays (31-32)

______________________ on weekends (33-34)

8. Do you rent or own your house?

(1) Rent ☐

(2) Own ☐ (35)____

9. How long have you lived in this house/apartment? __________ (36-38)

OBSERVATIONAL DATA

1. Date __________ (39-42)
day/month

2. Hour of Day __________ (43-44)

3. Weather:

Temperature __________ (45-46)

Fourths of Cloud Cover __________ (47)

Humidity __________ (48-49)

4. Building Construction:

a. Number of Stories in Building __________ (50)

b. Building material:

(1) Brick ☐

(2) Frame ☐

(3) Stucco ☐ (51)

(4) Asbestos Panels ☐

(5) Other (specify) ☐

c. Type of windows:

(1) Single ☐ (77

(2) Thermo-pane (double) ☐

d. Air conditioning:

(1) Central ☐ (3) none ☐ (78

(2) Window units ☐ (4) don't know ☐

5. Buffer features:

	Front	Back	Sides
Shrubs	52	53	54
Trees (Decid.)	55	56	57
Trees (Conif.)	58	59	60
Hedge	61	62	63
Wall/Solid Fence	64	65	66

[*Blank = 1, check = 2*] [*With reference to major noise source.*]

6. Type of dwelling unit:

(1) apartment ☐

(2) flat ☐

(3) row ☐ (67)

(4) semi-detached ☐

(5) detached ☐

7. If an apartment or flat, which floor? ________ (68-69)

[*When coding, col. (70) = 9*]

APPENDIX 2

VARIANCE AND SAMPLING DISTRIBUTION OF $\hat{EN}_p$

From equation (3.4) we have:

$$\hat{EN}_p = -\left(\frac{\hat{\alpha} + \log(\frac{1}{p} - 1)}{\hat{\beta}}\right) \qquad \text{(A1.1)}$$

So, $$\text{VAR}(\hat{EN}_p) = \text{VAR}(-\hat{EN}_p) = \text{VAR}\left(\frac{\hat{\alpha} + \log(\frac{1}{p} - 1)}{\hat{\beta}}\right)$$

Let $f = (\hat{\alpha} + \log(\frac{1}{p} - 1))/\hat{\beta}$ and let $\frac{\partial f}{\partial \alpha}$ denote $\left(\frac{\partial f}{\partial \hat{\alpha}}\bigg|\ \hat{\alpha} = \alpha\right)$ and similarly for $\hat{\beta}$; then

$$\text{VAR}\left(\frac{\hat{\alpha} + \log(\frac{1}{p} - 1)}{\hat{\beta}}\right) = \left(\frac{\partial f}{\partial \alpha}\right)^2 \text{VAR}(\hat{\alpha}) + 2\left(\frac{\partial f}{\partial \alpha}\right)\left(\frac{\partial f}{\partial \beta}\right) \text{COVAR}(\hat{\alpha}, \hat{\beta})$$
$$+ \left(\frac{\partial f}{\partial \beta}\right)^2 \text{VAR}(\hat{\beta})$$

(Bury, 1974), a consistent estimate of which is

$$\text{VAR}\left(\frac{\hat{\alpha} + \log(\frac{1}{p} - 1)}{\hat{\beta}}\right) = \frac{1}{\hat{\beta}}^2 \text{VAR}(\hat{\alpha}) + 2(\frac{1}{\hat{\beta}})\left(\frac{-\hat{\alpha} - \log(\frac{1}{p} - 1)}{\hat{\beta}^2}\right)$$
$$\cdot \text{COVAR}(\hat{\alpha}, \hat{\beta}) + \left(\frac{-\hat{\alpha} - \log(\frac{1}{p} - 1)}{\hat{\beta}^2}\right)^2 \text{VAR}(\hat{\beta})$$

$$= \frac{1}{\hat{\beta}^2}\left[\mathrm{VAR}(\hat{\alpha}) - 2\left(\frac{\hat{\alpha} + \log(\frac{1}{p} - 1)}{\hat{\beta}}\right)\mathrm{COVAR}(\hat{\alpha}, \hat{\beta}) + \left(\frac{\hat{\alpha} + \log(\frac{1}{p} - 1)}{\hat{\beta}}\right)^2 \mathrm{VAR}(\hat{\beta})\right] \qquad \text{(A1.2)}$$

Equation (A1.2) can be used to calculate the variance since all the terms in the equation will be known from the maximum likelihood estimation procedure (see Appendix 3).

With regard to the sampling distribution of $\hat{E}N_p$, the derivation which follows is for the special case where $\hat{\alpha}$ and $\hat{\beta}$ are independently distributed, i.e., where $\mathrm{COVAR}(\hat{\alpha}, \hat{\beta}) = 0$. The more general case of non-zero covariance is not treated since the purpose of the demonstration is merely to justify the failure in Chapter 5 to perform significance tests on the $\hat{E}N_p$, by showing that the probability density function (pdf) for $\hat{E}N_p$ does not correspond to any of the well known sampling distributions. Note that we need only consider the distribution of the term $\hat{\alpha}/\hat{\beta}$ since neither the constant term nor the negative sign in equation (A1.1) affect the result.

By the theory of maximum likelihood estimators $\hat{\alpha}$, $\hat{\beta}$ are asymptotically normally distributed with large sample means α, β and variances σ_1^2, σ_2^2 respectively. We wish to discover the pdf for $X = \hat{\alpha}/\hat{\beta}$. Assume that $\mathrm{COVAR}(\hat{\alpha}, \hat{\beta}) = 0$, then the joint pdf of $(\hat{\alpha}, \hat{\beta})$ is

$$f_{\hat{\alpha},\hat{\beta}}(\hat{\alpha}, \hat{\beta}) = \frac{1}{2\pi\sigma_1\sigma_2} \exp -\left[\frac{1}{2}\left\{\left(\frac{\hat{\alpha} - \alpha}{\sigma_1}\right)^2 + \left(\frac{\hat{\beta} - \beta}{\sigma_2}\right)^2\right\}\right]$$

Let $Y = \hat{\beta}$; then the transformation from $\hat{\alpha}$ and $\hat{\beta}$ to X and Y is

$$X = \hat{\alpha}/\hat{\beta}$$

$$Y = \hat{\beta}$$

with inverse transformation

$$\hat{\alpha} = XY$$

$$\hat{\beta} = Y$$

The determinant of the Jacobian of the inverse transformation is Y. Hence the joing pdf of X and Y is

$$f_{X,Y}(x, y) = f_{\hat{\alpha},\hat{\beta}}(xy, y) \cdot |y|$$

(DeGroot, 1975, 131-138; Giri, 1974, 88-97; Apostol, 1974, 421-422), and so the marginal pdf for X is given by

$$f_X(x) = \int_{-\infty}^{\infty} |y| \, f_{\hat{\alpha},\hat{\beta}}(xy, y) \, dy$$

$$= \int_{-\infty}^{\infty} |y| \, \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[-\frac{1}{2}\left\{\left(\frac{xy - \alpha}{\sigma_1}\right)^2 + \left(\frac{y - \beta}{\sigma_2}\right)^2\right\}\right] dy$$

$$= 2\int_0^\infty y \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[-\frac{1}{2}\left[\left\{\frac{xy-\alpha}{\sigma_1}\right\}^2 + \left\{\frac{y-\beta}{\sigma_2}\right\}^2\right]\right]dy \qquad \text{(A1.3)}$$

$$= \frac{1}{\pi\sigma_1\sigma_2} \int_0^\infty y \exp\left[-\frac{1}{2}\left[\left\{\frac{xy-\alpha}{\sigma_1}\right\}^2 + \left\{\frac{y-\beta}{\sigma_2}\right\}^2\right]\right]dy$$

The term $\left\{\frac{xy-\alpha}{\sigma_1}\right\}^2 + \left\{\frac{y-\beta}{\sigma_2}\right\}^2$ can be expanded to

$$y^2 \frac{x^2\sigma_2^2 + \sigma_1^2}{\sigma_1^2\sigma_2^2} - y \frac{2x\alpha\sigma_2^2 + 2\beta\sigma_1^2}{\sigma_1^2\sigma_2^2} + \frac{\alpha\sigma_2^2 + \beta^2\sigma_1^2}{\sigma_1^2\sigma_2^2}$$

So set

$$a = \frac{x^2\sigma_2^2 + \sigma_1^2}{2(\sigma_1^2\sigma_2^2)}$$

$$b = \frac{x\alpha\sigma_2^2 + \beta\sigma_1^2}{\sigma_1^2\sigma_2^2}$$

$$c = \frac{\alpha^2\sigma_2^2 + \beta^2\sigma_1^2}{2(\sigma_1^2\sigma_2^2)}$$

and substitute into (A1.3) to give

$$f_X(x) = \frac{1}{\pi\sigma_1\sigma_2} \int_0^\infty y \exp(-(ay^2 - by + c))\, dy$$

Let $u = (ay^2 - by + c)$

So, $\quad du = (2ay - b)\, dy$

And,

$$f_X(x) = \frac{1}{\pi\sigma_1\sigma_2}\left[\frac{1}{2a}\int_0^\infty e^{-u}\,du + \int_0^\infty \frac{b}{2a}\exp(-(ay^2 - by + c))\,dy\right]$$

$$= \frac{1}{\pi\sigma_1\sigma_2}\left[\frac{1}{2a} + \int_0^\infty \frac{b}{2a}\exp(-(ay^2 - by + c))\,dy\right]$$

Now, $ay^2 - by + c = a(y - \frac{b}{2a})^2 + c - \frac{b^2}{4a}$ (completing the square)

Let $k = c - \frac{b^2}{4a}$

So, $ay^2 - by + c = a(y - \frac{b}{2a})^2 + k$

and,

$$f_X(x) = \frac{1}{\pi\sigma_1\sigma_2}\left[\frac{1}{2a} + \frac{b}{2a}\int_0^\infty \exp(-\{a(y - \frac{b}{2a})^2 + k\})\,dy\right]$$

$$= \frac{1}{\pi\sigma_1\sigma_2}\left[\frac{1}{2a} + \frac{b}{2a}\int_0^\infty \exp(-(as^2 + k))\,ds\right]$$

(setting $s = y - \frac{b}{2a}$ so that $ds = dy$)

$$= \frac{1}{\pi\sigma_1\sigma_2}\left[\frac{1}{2a} + \frac{b}{2a}\int_0^\infty e^{-as^2}/e^k\,ds\right]$$

$$= \frac{1}{\pi\sigma_1\sigma_2}\left[\frac{1}{2a} + \frac{b}{2a}\cdot\frac{1}{e^k}\int_0^\infty e^{-as^2}\,ds\right]$$

But, $\int_0^\infty e^{-as^2}\,ds = \frac{1}{2\sqrt{a}}\sqrt{\pi}$

So,

$$f_X(x) = \frac{1}{\pi\sigma_1\sigma_2}\left[\frac{1}{2a} + \frac{b}{2a}\cdot\frac{1}{e^k}\left(\frac{\sqrt{\pi}}{2\sqrt{a}}\right)\right]$$

whence, substituting back in for a, b and k, expanding and simplifying gives:

$$f_X(x) = \frac{\sigma_1\sigma_2}{\pi(x^2\sigma_2^2 + \sigma_1^2)} + \frac{X\alpha\sigma_2^2 + \beta\sigma_1^2}{(x^2\sigma_2^2 + \sigma_1^2)^{3/2}} \cdot \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left\{\frac{(x\beta - \alpha)^2}{x^2\sigma_2^2 + \sigma_1^2}\right\}\right\} \qquad (A1.\ )$$

which is the desired result. The first term is something like a Cauchy density, the second something like a normal. Note that when the distributions of both $\hat{\alpha}$ and $\hat{\beta}$ are located at the origin (i.e. $\alpha, \beta = 0$), the second term disappears leaving

$$f_X(x) = \frac{\sigma_1\sigma_2}{\pi(x^2\sigma_2^2 + \sigma_1^2)}$$

which is the result given by Gnedenko (1967, 186-187) for the quotient of two independently normally distributed random variables located at the origin.

APPENDIX 3

ESTIMATION PROCEDURE

Let Y denote a dichotomous response variable which, for a particular individual i, assumes values 1 and 0 with probabilities P_i = Prob(Y_i=1) and Q_i = Prob(Y_i=0) respectively. For example Y might be a disturbance variable with Y_i = 1 if individual i is disturbed by noise. Let $\underline{\underline{Y}}$ = (Y_1, ..., Y_N) denote the actual response pattern of a sample of N individuals.

From equation (2.15) we can write

$$P_i = [1 + \exp(-\underline{\underline{\beta X_i}})]^{-1} \qquad (A3.1)$$

where $X_{1i} = 1$ $\qquad i = 1, \ldots, N$

so that β_1 represents the constant term α in equation (2.15). The vector of independent variables will also include the noise metric.

Calibration of the binary logit model given by (A3.1) involves estimating the parameter vector $\underline{\underline{\beta}} = (\beta_1, \ldots, \beta_K)$. The principle of maximum likelihood estimation used in the present study states that a consistent, asymptotically normal estimate of $\underline{\underline{\beta}}$ is given by that value $\hat{\underline{\underline{\beta}}}$ which maximizes the probability of the actual response pattern $\underline{\underline{Y}}$ actually occurring, that is which maximizes:

$$L = \prod_{i=1}^{N} P_i^{Y_i} Q_i^{1-Y_i}$$

Equivalently, but more conveniently, we can maximize:

$$\log L = \sum_{i=1}^{N} Y_i \log P_i + (1 - Y_i) \log Q_i \qquad \text{(A3.2)}$$

The quantity logL is known as the log likelihood of $\underline{\underline{Y}}$.

The value of the parameter vector $\underline{\underline{\beta}}$ which maximizes the log likelihood is that value at which:

$$\frac{\partial \log L}{\partial \beta_k} = 0 \qquad k = 1, \ldots, K \qquad \text{(A3.3)}$$

Equation (A3.3) is usually solved by interative numerical methods. The method which was used in the present study is that of Newton-Raphson (Kreyszig, 1972, 641-642). According to this method, if $\hat{\underline{\underline{\beta}}}_0$ is an initial estimate of the parameter vector $(\beta_1, \ldots, \beta_K)$, then improved estimates $\underline{\underline{\beta}}_1$ are obtained from

$$\underline{\underline{\beta}}_1 = \underline{\underline{\beta}}_0 - H_{\log L}(\underline{\underline{\beta}}_0)^{-1} (\nabla \log L_0) \qquad \text{(A3.4)}$$

where: $H_{\log L}(\underline{\underline{\beta}}_0)$ is the Hessian matrix of elements $\frac{\partial^2 \log L}{\partial \beta_k \partial \beta_j}$ evaluated at $\underline{\underline{\beta}}_0$

$\nabla \log L_0$ is the gradient of logL i.e. a column vector of elements $\frac{\partial \log L}{\partial \beta_k}$ for k = 1, ..., k.

(Pamik, 1976). In logit analysis the estimation procedure turns out to be very easy to computerize since

$$\frac{\partial \log L}{\partial \beta_k} = \sum_{i=1}^{N} \frac{Y_i}{P_i}\frac{\partial P_i}{\partial \beta_k} + \left(\frac{(1 - Y_i)}{Q_i}\right)\frac{\partial Q_i}{\partial \beta_k}$$

$$= \sum_{i=1}^{N} \frac{Y_i}{P_i}\frac{\partial P_i}{\partial \beta_k} - \left(\frac{1 - Y_i}{Q_i}\right)\frac{\partial P_i}{\partial \beta_k}$$

$$= \sum_{i=1}^{N} \frac{\partial P_i}{\partial \beta_k}\left(\frac{Y_i}{P_i} - \frac{1 - Y_i}{Q_i}\right)$$

$$= \sum_{i=1}^{N} x_{ik} P_i Q_i \left(\frac{Y_i - P_i}{P_i Q_i}\right)$$

$$= \sum_{i=1}^{N} x_{ik}(Y_i - P_i)$$

Also,

$$\frac{\partial^2 \log L}{\partial \beta_k \partial \beta_j} = \frac{\partial}{\partial \beta_k} \sum_{L=1}^{N} x_{ij}(Y_i - P_i)$$

$$= -\sum_{i=1}^{N} \frac{\partial}{\partial \beta_k}(x_{ij} P_i)$$

$$= -\sum_{i=1}^{N} x_{ij}\, x_{ik}\, P_i Q_i$$

APPENDIX 4

THE LOGIT PROGRAM

```
      PROGRAM MAIN(INPUT,OUTPUT,HARPS,TAPE3=INPUT,
     1TAPE2=OUTPUT,TAPE1=HARPS)
      COMMON/BETA/B(9)
      COMMON/L4/SCAL(9)
      COMMON/[illegible]/VNAMES(9)
      COMMON/LEVELS/Z(9)
C
C     BINARY LOGIT PROGRAM WRITTEN BY [illegible]
C     MCMASTER UNIVERSITY, DECEMBER 1977
C     PROGRAM READS IN PARAMETERS, PRINTS A DESCRIPTION OF THE
C     PROBLEM, INITIALIZES B, [illegible]
C
      READ(3,30)NVAR,NCASE,NOPT,NCONS,NSCALE,[illegible]
     1TOL,[illegible]
      READ(3,200) (VNAMES(J),J=1,NVAR)
      IF(NOPT.EQ.0) GO TO 10
      READ(3,40) (B(M),M=1,NVAR)
      GO TO 25
   10 DO 20 M=1,NVAR
      B(M)=0.001
   20 CONTINUE
   25 CONTINUE
      IF(NSCALE.EQ.0) GO TO 100
      READ (3,110) (SCAL(M),M=1,NVAR)
      GO TO 120
  100 DO 130 M=1,NVAR
      SCAL(M)=1.0
  130 CONTINUE
  120 CONTINUE
      READ(3,140) (Z(J),J=1,NEQ)
      CALL MANIP(NVAR,NCASE,NCONS,NSCALE,NCUCASE)
      WRITE(2,50) NCUCASE,NVAR,[illegible]
      IF(NOPT.EQ.1) WRITE(2,60)
      IF(NOPT.EQ.1) WRITE(2,70) (B(M),M=1,NVAR)
      IF(NSCALE.EQ.1)WRITE(2,80)
      IF(NSCALE.EQ.1)WRITE(2,90) (SCAL(M),M=1,NVAR)
      WRITE(2,210) (VNAMES(J),J=1,NVAR)
      CALL [illegible]
```

```
      CALL STATS(NVAR,NGDCASE,ITERS,NEQ)
   30 FORMAT(I1,I6,I11,3F6.0)
   40 FORMAT(2F5.0)
   50 FORMAT(1H1,45X,'DATA CONTAINS',I5,1X,'GOOD CASES'//
     145X,'MODEL IS ESTIMATED FOR',I2,1X,'VARIABLES'//
     145X,'32H',I2,1X,'CONSTANT(S)'//)
   60 FORMAT(/1H0,'INITIAL VALUES FOR B ARE')
   70 FORMAT(1H0,6(F15.9,3X))
   80 FORMAT(/1H0,'SCALING FACTORS FOR B ADJUSTMENT ARE')
   90 FORMAT(1H0,6(F5.1,3X))
  110 FORMAT(2F5.0)
  140 FORMAT(9F5.0)
  200 FORMAT(8A10)
  210 FORMAT(///' VARIABLES ANALYZED ARE (IN ORDER)',9A10)
      STOP
      END
      SUBROUTINE MANIP(K,N,NC,NS,NGOOD)
      COMMON/RESPONS/Y1(1786)
      COMMON/INDVARS/X(1786,6)
      COMMON/L4/SCAL(9)
      DIMENSION INTS(9)
C
C     ROUTINE READS IN RAW DATA ON REQUIRED VARIABLES
C     DELETES BAD (MISSING DATA) CASES, AND CREATES
C     THE NEW VARIABLES FOR ANALYSIS
C
      NBAD=0
      DO 10 I=1,N
      READ(1,100)IY,DN,LEVLD,IRAT,ISENSE,(INTS(J),J=1,9)
C
C     SKIP OVER CASES WITH MISSING DATA
C
      IF(IY.EQ.0) GO TO 50
      IF(DN.EQ.0) GO TO 50
      IF(ISENSE.EQ.0.OR.IRAT.EQ.0.OR.IRAT.EQ.6) GO TO 50
C
C     DEFINE INTERRUPTION VARIABLES
C
      NSLP=0
      IF(INTS(1).EQ.2.OR.INTS(4).EQ.2.OR.INTS(7).EQ.2) NSLP=1
      NRLX=0
      IF(INTS(2).EQ.2.OR.INTS(5).EQ.2.OR.INTS(8).EQ.2) NRLX=1
      NCON=0
      IF(INTS(3).EQ.2.OR.INTS(6).EQ.2.OR.INTS(7).EQ.2) NCON=1
C
C     DEFINE POPULATION SUBGROUP FOR ANALYSIS
C
      IF(ISENSE.LE.3.OR.IRAT.LE.2.OR.NRLX.EQ.0.OR.NSLP.EQ.1) GO TO 50
C
C
C     DEFINE DEPENDENT AND INDEPENDENT VARIABLES
C
      Y1(I-NBAD)=0.0
      IF(IY.GE.6) Y1(I-NBAD)=1.0
      IF(NC.EQ.0) GO TO 70
      X(I-NBAD,1)=1.0
   70 CONTINUE
```

```
      X(I-NBAD,1+NC)=PN
      GO TO 10
   50 NBAD=NBAD+1
   10 CONTINUE
      NGOOD=N-NBAD
   20 CONTINUE
      DO 60 I=1,NGOOD
      DO 95 M=1,K
      X(I,M)=X(I,M)/SCAL(M)
   95 CONTINUE
   60 CONTINUE
  100 FORMAT(10X,I1,12X,F4.0,3X,I1,2X,I1,8X,4I1,2(4),3I1))
      RETURN
      END
      SUBROUTINE ESTMATE(K,N,NC,NS,INT,TOL1,TOL2,TOL3)
      COMMON/RESPONS/ Y1(1786)
      COMMON/INDVARS/X(1786,6)
      COMMON/PROBS/P(1786)
      COMMON/INVERSE/WINV(9,9)
      COMMON/BETA/B(9)
      COMMON/L4/SCAL(9)
      COMMON/ONLY/ELSTARB
      DIMENSION EXPARG(1786),PQ(1786)
      DIMENSION W(9,9),DELTAB(9),D(9)
      DIMENSION ELSTAR(15)
C
C     BEGIN ITERATION HERE
C     FIRST INITIALIZE D,W,DELTAB,EXPARG
C
      ICON=0
      DO 115 L=1,15
      ELSTAR(L)=0.0
      DO 10 M=1,K
      DELTAB(M)=0.0
      D(M)=0.0
      DO 10 J=1,K
      W(M,J)=0.0
   10 CONTINUE
      DO 20 I=1,N
      EXPARG(I)=0.0
   20 CONTINUE
C
C     NOW CALCULATE THE PROBABILITIES OF RESPONSE
C
   25 DO 30 I=1,N
      DO 40 M=1,K
      EXPARG(I)=EXPARG(I)+B(M)*X(I,M)
   40 CONTINUE
      P(I)=1.0/(1.0+EXP(-EXPARG(I)))
      PQ(I)=P(I)*(1.0-P(I))
      PROB=1.0-P(I)
      IF(Y1(I).EQ.1.0) PROB=P(I)
      ELSTAR(L)=ELSTAR(L)+ALOG(PROB)
      ELSTARB=ELSTAR(L)
   30 CONTINUE
C
C
```

```
C       NOW DERIVE THE HESSIAN MATRIX W, AND THE VECTOR OF FIRST
C       PARTIALS D, BOTH EVALUATED AT THE CURRENT VALUE OF B
C
C
        DO 50 M=1,K
        DO 60 J=1,K
        DO 70 I=1,N
        W(M,J)=W(M,J)+PQ(I)*X(I,M)*X(I,J)
     70 CONTINUE
        W(M,J)= W(M,J)
        W(J,M)=W(M,J)
     60 CONTINUE
        DO 80 I=1,N
        D(M)=D(M)+X(I,M)*(Y1(I)-P(I))
     80 CONTINUE
     50 CONTINUE
C
C       NEXT INVERT MATRIX W, CALCULATE DELTAB
C       AND WRITE INTERMEDIATE OUTPUT
C
        CALL MI(W,?,K,DET,IDET,WINV)
        IF(JCON.EQ.1) GO TO 160
        DO 90 M=1,K
        DO 90 J=1,K
        DELTAB(M)=DELTAB(M)+WINV(M,J)*D(J)
     90 CONTINUE
        IF(INT.EQ.0) GO TO 500
        WRITE(2,550)L
        WRITE(2,560)
        WRITE(2,520) (B(M),M=1,K)
        WRITE(2,570) ELSTAR(L)
        WRITE(2,510)
        DO 600 M=1,K
        WRITE(2,520) (W(M,J),J=1,K)
    600 CONTINUE
        DO 610 M=1,K
        WRITE(2,520) (WINV(M,J),J=1,K)
    610 CONTINUE
        WRITE(2,530)
        WRITE(2,520) (D(M),M=1,K)
        WRITE(2,540)
        WRITE(2,520) (DELTAB(M),M=1,K)
C
C       TEST FOR CONVERGENCE, THEN INCREMENT B
C
    500 SUM=0.0
        ICONV1=1
        DO 650 M=1,K
        TEST1=ABS(DELTAB(M)/B(M))
        IF(TEST1.GT.TOL1) ICONV1=0
        SUM=SUM+(DELTAB(M)/B(M))**2
    650 CONTINUE
        AK=K
        ICONV2=0
        TEST2=SQRT(SUM/AK)
        IF(TEST2.LT.TOL2) ICONV2=1
        ICONV=??
```

```
      IF(L.EQ.1) GO TO 660
      TEST2=ABS((ELSTAR(L)-ELSTAR(L-1))/ELSTAR(L))
      IF(TEST2.LT.TOL3) ICONV3=1
  660 JCON=1
      IF((ICONV1.EQ.0.OR.ICONV2.EQ.0).AND.
     1ICONV3.EQ.0) JCON=0
      DO 620 M=1,K
      B(M)=B(M)-DELTAB(M)
  620 CONTINUE
  115 CONTINUE
C
C     READJUST B AND THE COVARIANCES, THEN
C     WRITE B AND THE VARIANCE/COVARIANCE MATRIX
C
  160 WRITE(2,630)
      WRITE(2,640) ICONV1,ICONV2,ICONV3
      WRITE(2,400)
      DO 170 M=1,K
      DO 180 J=1,K
      WINV(M,J)=-WINV(M,J)*(1.0/SCAL(M))*(1.0/SCAL(J))
  180 CONTINUE
      B(M)=B(M)/SCAL(M)
      WRITE(2,410) (WINV(M,J),J=1,K)
  170 CONTINUE
      WRITE(2,420)
      WRITE(2,430) (B(M),M=1,K)
  400 FORMAT(//////" COVARIANCE MATRIX FOR B")
  410 FORMAT(1H ,6(F15.9,3X))
  420 FORMAT(///1H ,"OPTIMUM VALUES OF B-COEFFICIENTS"/)
  430 FORMAT(1H ,6(F15.9,3X))
  510 FORMAT(1H0,"W AND WINV ARE RESPECTIVELY")
  520 FORMAT(1H0,6(F15.7,3X))
  530 FORMAT(1H0,"THE VECTOR OF FIRST PARTIALS")
  540 FORMAT(1H0,"CHANGES IN B FROM LAST ITERATION")
  550 FORMAT(////1H0,"ITERATION NUMBER",I2)
  560 FORMAT(1H0,"VALUES FOR B AT START OF ITERATION ARE")
  570 FORMAT(1H0,"LIKELIHOOD =",F15.3)
  630 FORMAT(//////" CONVERGENCE PERFORMANCE -- A ZERO"/
     1" INDICATES TEST NOT SATISFIED")
  640 FORMAT(/" TEST 1     ",I1//" TEST 2     ",I1/
     1/" TEST 3     ",I1)
      RETURN
      END
      SUBROUTINE STATS(K,N,NC,NE)
      COMMON/RESPONS/ Y1(1786)
      COMMON/PROBS/P(1786)
      COMMON/INVERSE/WINV(9,9)
      COMMON/BETA/B(9)
      COMMON/ONLY/ELSTARB
      DIMENSION T(9),VAR(9)
      DIMENSION ED(9),VARED(9),SED(9)
      COMMON/LEVELS/Z(9)
      DO 10 M=1,K
      VAR(M)=WINV(M,M)
      T(M)=ABS(B(M)/SQRT(VAR(M)))
   10 CONTINUE
      WRITE(2,70)
```

```
      WRITE(2,40)K
      WRITE(2,50) (T(M),M=1,K)
C
C     NOW COMPUTE LIKELIHOOD RATIO STATISTICS.
C
      R=0.0
      S=0.0
      DO 20 I=1,N
      R=R+Y1(I)
      Y2 =0.0
      IF(P(I).GE.0.5) Y2=1.0
      S=S+ABS(Y1(I)-Y2)
   20 CONTINUE
      RN=N
      ELSTARO=RN*ALOG(0.5)
      ELSTARC=R*ALOG(R/RN)+(RN-R)*ALOG((RN-R)/RN)
      RHOSQRO=1.0-(ELSTARB/ELSTARO)
      RHOSQRC=1.0-(ELSTARB/ELSTARC)
      CHISQRO=2.0*(ELSTARB-ELSTARO)
      CHISQRC=2.0*(ELSTARB-ELSTARC)
      PERCOR=((RN-S)/RN)*100.0
      PERNO=((RN-R)/RN)*100.0
      PERRESP=(R/RN)*100.0
      WRITE(2,60)ELSTARB
      WRITE(2,70)RHOSQRO
      WRITE(2,80)RHOSQRC
      WRITE(2,90)CHISQRO,K
      WRITE(2,100)CHISQRC,K-NC
      WRITE(2,110) PERCOR,PERNO,PERRESP
C
C     FINALLY THE EFFECTIVE DOSE LEVELS
C
      IF(NE.EQ.0) GO TO 120
      WRITE(2,150)
      T1= B(1)/B(2)
      DO 140 J=1,NE
      T2 = ALOG((1.0/Z(J))-1.0)/B(2)
      ED(J)=-T1-T2
      VARED(J)=(1.0/B(2))**2.0*(VAR(1)-2.0*(T1+T2)*UNU(1,2)+
     1((T1+T2)**2.0)*VAR(2))
      SED(J)=SQRT(VARED(J))
      WRITE(2,160) Z(J),ED(J),VARED(J),SED(J)
  140 CONTINUE
  120 CONTINUE
   30 FORMAT(///,45X,'GOODNESS OF FIT AND T-STATISTICS'///)
   40 FORMAT(1H ,'T-VALUES FOR B(M),M=1-',I2/)
   50 FORMAT(1H ,6(F8.4,3X)//)
   60 FORMAT(///' L*(B)-MAXIMUM LIKELIHOOD=',F15.3//)
   70 FORMAT(1H ,'RHOSQUARE(O) =',F12.9/)
   80 FORMAT(1H ,'RHOSQUARE(C) =',F12.9//)
   90 FORMAT(1H ,'CHI-SQUARE(O)=',F15.3,3X,'WITH',I2,1X,
     1'DEGREES OF FREEDOM')
  100 FORMAT(/1H ,'CHI-SQUARE(C)=',F15.3,3X,'WITH',I2,1X,
     1'DEGREES OF FREEDOM')
  110 FORMAT(//1H ,'PERCENTAGE OF RESPONSES CORRECTLY PREDICTED',
     1F6.2,' IN A SAMPLE WITH A ',F5.1,'/',F5.1,' SPLIT')
  150 FORMAT(///1H ,'PROBABILITY', [illegible]
     18X,'STANDARD ERROR')
  160 FORMAT(/,6X,F6.3,14X,F6.2,2(10X,F8.2))
      RETURN
      END
```

BIBLIOGRAPHY

Apostol, T.M. (1974) *Mathematical Analysis*. Reading, Mass.: Addison-Wesley Publishing Company.

Ashton, W.D. (1972) *The Logit Transformation*. London: Charles Griffin and Company Limited.

Bauer, R.A. (1970) "Predicting the Future". In J.D. Chalupnik (ed.), *Transportation Noises*, Seattle: University of Washington Press, 245-256.

Berglund, B., Berglund, U. and Lindvall, T. (1974) "Scaling of Annoyance in Epidemiological Studies". Paper presented at the CEC-WHP-EPA International Symposium "Recent Advances in the Assessment of the Health Effects of Environmental Pollution", Paris, June 1977.

Borsky, P.N. (1970) "The Use of Social Surveys for Measuring Community Responses to Noise Environments". In J.D. Chalupnik (ed.), *Transportation Noises*, Seattle: University of Washington Press, 219-227.

Brummell, A.C. (1977) A Theory of Intraurban Residential Mobility Behaviour. Unpublished Ph.D. dissertation, Department of Geography, McMaster University, Hamilton, Ontario.

Bury, K.V. (1975) *Statistical Models in Applied Science*. New York: John Wiley and Sons Inc.

Cox, D.R. (1970) *The Analysis of Binary Data*. London: Methuen and Co. Ltd.

DeGroot, M.H. (1975) *Probability and Statistics*. Reading, Massachusetts: Addison-Wesley Publishing Company.

Domencich, T.A. and McFadden, D. (1975) *Urban Travel Demand: A Behavioral Analysis*. Amsterdam: North Hollan/American Elsevier.

Finney, D.J. (1971) *Probit Analysis*. Third Edition, Cambridge: C.U.P.

Fisher, R.A. (1922) "On the Mathematical Foundations of Theoretical statistics". In W.A. Shewart (ed.), *Contributions to Mathematical Statistics*, New York: John Wiley and Sons Inc., 1950.

Giri, N.C. (1974) *Introduction to Probability and Statistics*. New York: Marcel Dekker, Inc.

Gnedenko, B.V. (1967) *The Theory of Probability*. Fourth Edition, New York: Chelsea Publishing Company, Russian Translation by B.C. Seckler.

Goldstein, A. (1964) *Biostatistics: An Introductory Text*. New York: Macmillan.

Griffiths, I.D. and Langdon, F.J. (1968) "Subjective Response to Road Traffic Noise". *Journal of Sound and Vibration*, 8(1), 16-32.

Gurland, J., Lee, I. and Dahm, P.A. (1960) "Polychotomous Quantal Response in Biological Assay". *Biometrics*, 16, 382-397.

Hall, F.L., Taylor, S.M. and Birnie, S.E. (1977) "Community Response to Road Traffic Noise". Report submitted to the Ford Motor Company, December 1977.

Hall, F.L., Birnie, S. and Taylor, S.M. (1978) "The Effectiveness of Shielding in Reducing the Adverse Impacts of Highway Traffic Noise". Paper presented at the 57th Annual Meeting of the TRB, Washington, D.C., January 1978.

Hazard, W.R. (1971) "Predictions of Noise Disturbance Near Large Airports". *Journal of Sound and Vibration*, 15, 425-445.

Kreyszig, E. (1972) *Advanced Engineering Mathematics*. New York: John Wiley and Sons, Inc.

Lamure, C. and Bacelon, M. (1976) "La Gene due au Bruit de la Circulation Automobile: Une Enquete aupres de riverains d'autoroutes". Report No. 88, Cahier 762, Centre Scientifique et Technique du Batiment, Paris, October 1976.

Langdon, F.J. (1976) "Noise Nuisance Caused by Road Traffic in Residential Areas: Part I". *Journal of Sound and Vibration*, 47, 243-263.

Langdon, F.J. and Buller, I.B. (1977) "Road Traffic Noise and Disturbance to Sleep". *Journal of Sound and Vibration*, 50, 13-28.

Mantel, N. (1966) "Models for Complex Contingency Tables and Polychotomous Dosage Response Curves". *Biometrics*, 22, 83-95.

McKennell, A.C. (1970) "Noise Complaints and Community Action". In J.D. Chalupnik (ed.), *Transportation Noises*, Seattle: University of Washington Press, 228-244.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K. and Bert, D.H. (eds.) (1975) *Statistical Package for the Social Sciences*. Second Edition, New York: McGraw-Hill.

Panik, M.J. (1976) *Classical Optimization: Foundations and Extensions*. Amsterdam: North Holland/American Elsevier.

Quandt, R.E. and Baumol, W.J. (1966) "The Demand for Abstract Transport Modes: Theory and Measurement". *Journal of Regional Science*, 6(2), 13-26.

Rylander, R., Sorensen, S. and Kajland, A. (1976) "Traffic Noise Exposure and Annoyance Reactions". *Journal of Sound and Vibration*, 47, 237-242.

Schultz, T.J. (1972) *Community Noise Ratings*. London: Applied Science Publishers Ltd.

Schultz, T.J. (1977) "Synthesis of Social Surveys on Noise Annoyance". *Journal of the Acoustical Society of America*, forthcoming.

Starkie, D.N.M. and Johnson, D.M. (1975) *The Economic Value of Peace and Quiet*. Farnborough, Harts.: Saxon House.

Tardiff, T.J. (1976) "A Note on Goodness-of-Fit Statistics for Probit and Logit Models". *Transportation*, 5, 377-388.

Taylor, S.M. and Hall, F.L. (1977a) "Factors Affecting Response to Road Noise". *Environment and Planning* A, 9, 585-597.

Taylor, S.M. and Hall, F.L. (1977b) "Impact of Road Traffic Noise on Residential Areas". Resport to the Canada Council, Project S76-0462.

Taylor, S.M., Gertler, M. and Hall, F.L. (1978) "Regulatory Implication of Individual Reactions to Road Traffic Noise". Paper presented at the 57th Annual Meeting of the TRB, Washington, January 1978.

Tulloch, J.R. and MacMillan, J.R. (1973) "A Micro-Analytic Model of Migration Behaviour". *Regional Science Perspectives*, 3, 105-126.

United States Environmental Protection Agency (1974) *Information on Levels of Environmental Noise Requisite to Protect Public Health and Welfare with an Adequate Margin of Safety*. Washington, D.C.: U.S. Government Printing Office.

Wilson, A.G. (1974) *Urban and Regional Models in Geography and Planning*. London: John Wiley and Sons.

Wrigley, N. (1976) *Introduction to the Use of Logit Models in Geography*. Concepts and Techniques in Modern Geography, 10, Norwich: Geo Abstracts.