Working Paper No. 337

MULTIPLE SITE DEMAND MODELS. PART II:
REVIEW OF EXISTING MODELS AND DEVELOPMENT OF NEW MODELS

W. Michael Hanemann

This is a slight revision of Chapter 9 of Nancy E. Bockstael, W. Michael Hanemann, and Ivar E. Strand (eds.), Measuring the Benefits of Water Quality Improvement Using Recreation Demand Models, Department of Agricultural and Resource Economics, University of Maryland, October, 1984.
CHAPTER 9

MULTIPLE SITE DEMAND:

A REVIEW OF EXISTING MODELS AND THE DEVELOPMENT OF NEW MODELS

Some Existing Demand Models

We start with the studies by Burt and Brewer (1971) and Cicchetti et al. (1976), both of which treat quality implicitly by estimating separate demand for each group of recreation sites, and both of which employ the linear demand system (1.24) in Chapter 8—treating this as an incomplete demand system in which the prices of other goods (i.e., q2, ..., qM) are subsumed in the coefficients (i.e., the a's or the b's). Although it appears that each data set contains instances of corner solutions, this fact
is ignored in the formulation and estimation of the demand systems. Unlike Burt and Brewer, Cicchetti et al., assume no income effects and impose the symmetry conditions across equations. Both versions of (24) are, in principle, consistent with the hypothesis of utility maximization and the underlying direct utility functions have been obtained by LaFrance and Hanemann (1984). However, the version with non-zero income effects can only satisfy the integrability conditions (6) if the commodities (visits to recreation sites) are assumed to be perfect complements, consumed in fixed proportions.

The other demand studies that we shall discuss introduce site quality characteristics explicitly, in the sense that they imply a utility function of the form \( u(x, b, z) \). An example is the gravity model employed by Wennnergren and Nielsen (1970) in which there is a single quality dimension, site capacity, and

\[
(1) \quad x \propto \left( \frac{b}{p} \right). \left( \sum_{i=1}^{N} \frac{b_{i}}{p_{i}} \right) \quad i=1, \ldots, N.
\]

Gravity models have subsequently been employed by a variety of authors, two particularly sophisticated examples being the
studies by Cesario and Knetsch (1976) and Sutherland (1982a). In Cesario and Knetsch, the demand system is

\[ x_i = f(y, a) \psi_i e^{(\sum \psi_j e_j)} , \quad 0 < \gamma < 1, \quad i = 1, \ldots, N, \]

where \( a \) is a vector of individual characteristics and \( \psi_1, \ldots, \psi_N \) are constructed indices of site quality. As long as \( \gamma < 1 \), this formulation implies that a decrease in the cost of visiting a site and/or an increase in its quality index have two effects: not only is existing recreation actually diverted away from other sites to the site in question, but also some new recreation activity is generated (i.e. \( \sum x_j \) increases). (If \( \gamma = 1 \), the latter effect vanishes and \( \sum x_j \) remains constant.

Sutherland's model actually involves four distinct components. In order to explain them, it is necessary to introduce a subscript for consumers, which has so far been suppressed. Let \( x_{it} \) be the demand for site \( i \) by consumer \( t \), \( P_{it} \) the cost to consumer \( t \) of visiting site \( i \), \( y_t \) his income, and \( a_t \) his other attributes (e.g. age, sex). Thus, in terms of our previous discussion, the utility function becomes \( u(x_i, y_t, a_t) \) and the demand functions for sites are

\[ x_{it} = h^i(P_{1t}, \ldots, P_{Nt}, b_1, \ldots, b_n, q_t, y_t; a_t) \quad i = 1, \ldots, N. \]

A key element in the demand model is an impedance function \( \theta_{it} = \theta(P_{it}) \) which is constructed as follows. Sutherland constructs the empirical probability density of the subset of \( P_{it} 's \) for which \( x_{it} > 0 \) and then sets \( \theta(\cdot) \) equal to a smoothed version of this empirical density. A feature of the function \( \theta(\cdot) \)
therefore, is that it is increasing over a (usually small) part of its domain, and decreasing over the remainder. Let $x_t = \sum x_{it}$ be the consumer total visitation of all sites, and $x_i = \sum x_{it}$ the total attendance of site $i$ by all consumers. The first two components of Sutherland's model are a trip production equation (i.e. a participation intensity equation)

$$(4) \quad x_t = \sum x_{it} = f(\sum b_i \theta(P_{it}), y_{it}; a)$$

and a site attractiveness model (i.e. an aggregate demand function for each site)

$$(5) \quad x_i = \sum x_{it} = \alpha b_i \theta(P_{it}) \gamma, \alpha, \beta, \gamma > 0 \quad i=1, \ldots, N$$

where $b_i$ is a measure of facilities available at site $i$, $\sum b_i \theta(P_{it})$ is a measure of overall availability of recreation opportunities to consumer $t$, and $\sum \theta(P_{it}) (\sum x_{it})$ is intended as a measure of the overall accessibility of site $i$ to the population of consumers. Equations (4) and (5) are each estimated by OLS. Using the fitted regression equations the predicted values $\hat{x}_1, \ldots, \hat{x}_T$ are obtained from (4) and the predicted values $\hat{x}_1, \ldots, \hat{x}_N$ are obtained from (5); these are substituted into the following gravity model

$$(6) \quad \hat{x}_{it} = \hat{x}_t \left( \frac{\theta(P_{it}) \hat{x}_i}{\sum \theta(P_{jt}) \hat{x}_j} \right)$$

to obtain predictions, $\hat{x}_{it}$, of each consumer's visit to each site. Finally, for the purpose of valuing each site, the predicted site visits, $\hat{x}_{it}$, are regressed on the site costs via an equation of the form

$$(7) \quad \ln(\hat{x}_{it}) = \delta_i + \lambda_i P_{it}$$

from which the Marshallian triangle is approximated.
By contrast with (3), the model (4) - (7) may appear to be overfitting a demand system. Moreover, as with (1) and (2), the demand model does not appear to be desirable from a utility maximization standpoint, nor does it make any particular allowance for the appearance of corner solutions which certainly abound in the data set. This is partially rectified in the next group of models we consider, which we call share models.

**Share Models**

It is convenient here (although be no means essential) to employ the separability assumption described in Chapter 8 and to work with the system of partial demand functions, modified so that they now include site quality indices. That is, the utility function is assumed to be \( u(x,b,z) = f(\bar{u}(x,b),z) \) and the partial demand system, derived from \( \bar{u}(\cdot) \), is

\[
(8) \quad x_i^* = h_i^x(p,b,y_x) \quad i = 1, \ldots, N
\]

where

\[
(9) \quad y_x = \Sigma p_i x_i = p_i h_i^x(p,b,y_x).
\]

In addition to summability, we assume that demand functions \( h_1^x(\cdot), \ldots, h_N^x(\cdot) \) possess the requisite properties of symmetry, negative semi-definiteness and, in particular, non-negativity

\[
(10) \quad h_i^x(p,b,y_x) \geq 0 \quad i = 1, \ldots, N.
\]

Suppose that we are interested in the share of recreation activity allocated to each site; we could either value shares, \( w_i = p_i x_i / y_x \), and value share functions.
or quantity shares, \( s_i = x_i / x_s \), where \( x_s = \sum_{1 \leq j \leq N} x_j \), and quantity share functions

\[
(12) \quad s_i = s_i(p,b,y) = \frac{h_i(p,b,y)}{\sum_{1 \leq j \leq N} h_j(p,b,y)} \quad i = 1, \ldots, N.
\]

Note that the latter allocate total visitation, \( x_s \), among individual sites as function of total expenditure on all sites, \( y_x \), and not as a function of \( x_s \) itself. Moreover, if the sub-utility function \( \bar{u}(x,b) \) happens to be homothetic in \( x \), the partial demand functions take the special form

\[
(13) \quad x_i = h_i(p,b,y_x) = \phi_i(p,b)y_x \quad i = 1, \ldots, N
\]

for some set of functions \( \phi_1(\cdot), \ldots, \phi_N(\cdot) \) which are each homogeneous of degree minus one in \( p \). In this case, the share equations become

\[
(14) \quad w_i = w_i(p,b,y) = \frac{\xi_i(p,b)}{\sum_{1 \leq j \leq N} \xi_j(p,b)} \quad i = 1, \ldots, N,
\]

where \( \xi_i(p,b) = p_i \phi_i(p,b) \), and

\[
(15) \quad s_i = s_i(p,b,y) = \frac{\phi_i(p,b)}{\sum_{1 \leq j \leq N} \phi_j(p,b)} \quad i = 1, \ldots, N.
\]

i.e. the share equations are independent of total expenditure on site recreation activity, \( y_x \). Regardless of whether \( \bar{u}(\cdot) \) is homothetic, the point we wish to emphasize is that modelling approaches based on the systems of share equations are entirely equivalent to modelling approaches based on the system of partial demand equations. Any system of share equations implies a corresponding partial demand system, and conversely; both systems
convey the same amount of information about consumer preferences and behavior.  

The situation changes, however, as soon as we introduce stochastic elements and begin to think in terms of statistical models. Depending on the stochastic specification, it could make a considerable difference whether we choose to estimate share or demand systems. For example, suppose that we introduce additive, normally distributed disturbance terms into the partial demand system (8). In order to accommodate the summability restriction (9), which induces a dependence among the disturbance terms, we actually assume that a subset of (N-1) of the x_i's have an (N-1)-dimensional multivariate normal distribution with mean vector \( h^X_1(p,b,y_x), ..., h^X_{N-1}(p,b,y_x) \) and some covariance matrix \( \Omega \) which is an \((N-1)\times(N-1)\) positive definite matrix. It is understood that the remaining consumption level is obtained via

\[
x_N = \frac{y_x - \sum_{i=1}^{N-1} p_i x_i}{p_N}
\]

- i.e. and hence is also normally distributed.

It should be evident that it matters greatly whether we estimate the partial demand system or the share system because, whereas the x_i's are multivariate normal, the distribution of the s_i's (the composition of dependent normal variates) is extremely complex and does not possess a closed form expression. (See Yatchew, 1983, for a discussion of the computational problems in estimating this distribution.) Conversely, suppose we assume that the shares are multivariate normal, with means given by \( s_i(p,b,y_x) \) and some covariance matrix \( \Sigma \). While this paves the way for direct estimation of the
share equations, it rules out estimation of the partial demand system because there is no closed form expression for the distribution of the $x_i$'s.

In these two examples one has to make a direct choice between a tractable distribution for the observed $x_i$'s and a tractable distribution for the observed $s_i$'s. When distribution other than the multivariate normal are used, however, this dilemma can sometimes be avoided. For example, if we assume that $x_1, \ldots, x_N$ are independently distributed gamma variates, the gamma parameter being $y_i = \delta_i(p,b,y_x) \quad i = 1, \ldots, N$, the $(N-1)$ shares $s_1, \ldots, s_{N-1}$ have the Dirichlet distribution with density

$$f(s_1, \ldots, s_{N-1}) = \frac{1}{N^\gamma} \prod_{i=1}^{N-1} s_i^{\gamma-1} (1-s_1-\cdots-s_{N-1})^{\gamma-1},$$

where it is understood that $s_N$ is obtained from the relation:

$$s_N = 1 - s_1 - \cdots - s_{N-1}.$$ 

This approach to the modelling of shares was proposed by Woodland (1979). The estimates of the $\gamma_i$'s, i.e. the estimation of the coefficients of $\delta_i(p,b,y_x)$, can be based on either the observed quantities demanded, $x_1, \ldots, x_N$ or the observed shares, $s_1, \ldots, s_N$. Similarly, suppose that the $x_i$'s have a N-dimensional multivariate lognormal distribution with parameters $\mu$ and $\Omega$, where

$$\mu_i = \delta_i(p,b,y) \quad i = 1, \ldots, N$$
and $\Omega$ is some NxN positive definite matrix, i.e. $(\ln x_1, \ldots, \ln x_N)$ is $N(\mu, \Omega)$. The $(N-1)$ shares $s_1, \ldots, s_{N-1}$ have Aitchison and Shen's (1980) logistic normal distribution with density

$$f(s_1, \ldots, s_{N-1}) = \frac{1}{2\pi^{N-1}} \left( \prod_{i=1}^{N-1} s_i \right) \left( 1 - \sum_{i=1}^{N-1} s_i \right)^{-1} \exp \left[ -\frac{1}{2} \left( (a-\xi)^T \Sigma (a-\xi) \right) \right].$$
where it is again understood that \( s_N = 1 - s_2 - \ldots - s_{N-1} \).

Here, \( a' = (a_1, \ldots, a_N) \),  
\( a_i \equiv \ln(s_i/s_N), i = 1, \ldots, N-1 \), \( \xi \equiv A\mathbf{u} \), and \( \Sigma = A\Lambda A' \), \( A \) being the \((N-1)\times N\) matrix

\[
A = [I_{N-1}, -e_{N-1}]
\]

where \( I_{N-1} \) is the identity matrix and \( e_{N-1} \) a vector of \((N-1)\)'s.

In this case, too, the vector \( (i.e., \text{the functions} \)

\[
R^{X}_i(p, b, y_x) \quad i = 1, \ldots, N
\]

and the matrix \( \Omega \) can be estimated equally well from the observed demands or the observed shares.

Apart from considerations of ease of estimation, these four statistical models—normal demands, normal shares, gamma demands/Dirichlet shares, and lognormal demands/logistic normal shares—have different economic implications which are not unimportant. Both of the normal models can readily be employed when the data contain zero values for \( x_i \) and \( s_i \); the problem is that, even if the estimated \( R^{X}_i(\cdot) \) or \( s_i(\cdot) \) functions satisfy the nonnegativity requirement (which is not guaranteed to happen), these models imply that one can have negative values of \( x_i \) or \( s_i \) with some probability, which is a mis-specification from an economic point of view. To be sure, this is less likely to be a serious problem for the normal demands model since the density in the negative orthant is likely to be negligible, as Woodland (1979, p. 362) points out. The gamma/Dirichlet model ensures that the estimated \( R^{X}_i(\cdot) \) and \( s_i(\cdot) \) functions are positive, and it rules out the possibility of negative values for \( x_i \) and \( s_i \). It can be applied when the data contain zero values for \( x_i \) and \( s_i \) but only if \( 0 < x_i = R^{X}_i(p, b, y_x) \leq 1 \) i.e., the expected demand for the good does not exceed one unit. In this case the mode of the
the distribution of $\xi_i = 0$ and the gamma density does not possess the standard bell shape. If $\gamma_i = h_i^X(p,b,y) > 1$ the density possesses a conventional shape, but the domain over which it is defined excludes $x_i = 0$ or $s_i = 0$. The lognormal/logistic normal model does not impose any a priori restrictions on the sign or magnitude of $\mu_i = h_i^X(p,b,y,x)$, and it rules out the possibility of negative values for $x_i$ and $s_i$, but its domain is restricted to $x_i > 0$ and $0 < s_i < 1 \ i = 1,\ldots,N$. Thus it, too, cannot be applied to data containing zero values for the $x_i$'s or $s_i$'s.

Given the availability of observations on the demands as well as the shares, one might wonder why anybody would bother to estimate share systems. One reason why share systems have been employed in these circumstances (e.g. Morey (1981, 1984)) has to do with a factor that has not so far been mentioned and is not incorporated into any of the statistical models described above: the possibility that, since commodities may be indivisible, the $x_i$'s may be required to be non-negative integers.

Before describing Morey's model, we would like to make three observations on the general issue of integer-valued consumption levels. First, the extent to which this is a factor that ought to be accommodated in the estimation procedure is clearly a matter of degree, and depends on the particular data involved. If, as sometimes happens with recreational data, individuals consume very few units of the goods that they do buy (i.e., make very few
visits to sites), consumption demand ought perhaps to be modelled as integers. Second, integer-valued consumption levels can be modelled via demand systems without necessarily resorting to the multinomial distribution, for example by assuming that the $s_i$'s are independent Poisson variables. Third, aside from its implications for stochastic specification, the fact that commodities are indivisible and may be consumed only in integer units causes profound problems for the economic analysis of demand functions. This is because the consumer's utility maximization becomes an integer programming problem whose solution is very different from the smooth demand functions in (8) that are based on the presumption of an interior (i.e. non-integer-valued) solution. To assume that the $x_i$'s are Poisson (or multinomial) variates whose means are given by $\bar{H}_i^x(\cdot), \ldots, \bar{H}_N^x(\cdot)$ may involve an inconsistency between the economic and statistical specifications of the model.

1. The Morey Model

In employing the multinomial distribution, one is treating the $x_i$'s as "counts" of the occurrence of discrete events. The standard scenario underlying the multinomial distribution is that $R$ independent trials are held and, on each trial, $N$ mutually exclusive outcomes may occur, with $\pi_i$ being the probability of the $i$th outcome where $\pi_i > 0$ and $\sum_{i=1}^{N} \pi_i = 1$. Let $t_i$ be the number of times that the $i$th outcome occurs in $R$ trials. The probability of any outcome vector $(t_1, \ldots, t_N)$ is

\[
(18) \quad f(t_1, \ldots, t_N) = \frac{R!}{\prod_{j=1}^{N} t_j} \prod_{j=1}^{N} t_j \pi_j^{t_j}
\]

where it is understood that $t_j \geq 0$ for all $j$. 

9-11
\[ \sum_{j=1}^{N} t_j = R, \text{i.e.,} \ t_N = R - \sum_{j=1}^{N-1} t_j. \] Conditional on \( R \), some of the moments are

\begin{align*}
(19a) \quad & E\{t_i|R\} = R\pi_i \\
(19b) \quad & \text{var}\{t_i|R\} = R\pi_i(1-\pi_i) \\
(19c) \quad & \text{cov}\{t_i,t_j|R\} = -R\pi_i\pi_j.
\end{align*}

Following Morey, we equate the count \( t_i \) with the observed demand for the \( i \)th good \( x_i \), \( \pi_i \) with the share function \( s_i(p,b,y_x) \) in (12), and \( R \) with \( x. = \sum h^x_i(p,b,y_x) \), and write the density of the observed demands as

\[ f(x_1,\ldots,x_N) = \frac{(x_.)!}{N!} \pi_{x_1} \prod_{j=1}^{N} s_j(p,b,y_x)^{x_j}. \]

This density requires that \( s_i(p,b,y_x) > 0 \), and hence that \( h^x_i(p,b,y_x) > 0 \), \( i = 1,\ldots,N \), and its domain is \( \{x|x_i = 0,1,2,\ldots,x. , \sum x_i = x. \} \). Thus, it can readily be applied to data sets containing zero values of the observed \( x_i \)’s.

However, there is a conceptual problem in this application of the multinomial distribution to consumption data which needs to be recognized. The logic of the statistical model (18) is that the number of trials, \( R \), is exogenous and, therefore, this parameter may be ignored in maximizing (18) to obtain estimates of \( \pi_1,\ldots,\pi_N \).

In the consumption case, by contrast, the parameter \( x. \) is endogenous since it is simply the total consumption of all goods. Thus \( x. = \sum h^x_j(p,b,y_x) \) contains information on the coefficients to be estimated, and can hardly be ignored in maximizing the likelihood function derived from (20). Although this casts doubt on the theoretical suitability of (20) as a density for consumption data (regardless of whether consumption levels are integer-valued), this stochastic model may have some merit as a practical approximation.
When applying the model to recreation data, Morey uses two
different utility functions for generating the share equations.
In Morey (1981) he uses a CES subfunction

\[(21a) \quad u(x,b) = \left( \sum_{j=1}^{N} \psi_j(b_j)x_j^\rho \right)^{1/\rho} \quad \rho \leq 1, \quad \rho \neq 0\]

where \( \psi_j(b_j) \) is the overall quality index for site \( j \), and the
parameters to be estimated are \( \rho \) and the coefficients of the
\( \psi_j(*)'s \). This utility function implies homothetic partial
demand functions and, therefore, share functions which are
independent of \( y_x \) - see (15) above:

\[(21b) \quad s_i(p,b,y_x) = \left[ \psi_i(b_i)/p_i \right]^{\sigma} \left\{ \sum \psi_j(b_j)/p_j \right\}^{-1} \quad i = 1, \ldots, N\]

where \( \sigma = (1 - \rho)^{-1} \geq 0 \) is the (common) elasticity of substitution
among different commodities (recreation sites). However, because
of the homotheticity, this utility model implies that all
commodities have a unitary income elasticity of demand, which is
implausible in the recreation context. Morey (1984) recognizes
this and employs instead the following version of Pollak and
Wales' (1978) Quadratic Expenditure System indirect utility
function

\[(22a) \quad \tilde{v}(p,b,y_x) = -\frac{g(p,b)}{y_x} + \frac{q(p,b)}{f(p,b)}\]

where

\[g(p,b) = \sum \psi_j(b_j) p_j^{1-\sigma} \]

and

\[f(p,b) = \sum \psi_j(b_j) p_j^{1-\sigma}\]
\( \psi_j(b_j) \) and \( \phi_j(b_j) \) being different quality indices for site \( j \). The parameters to be estimated are now \( \sigma \) and the coefficients of the \( \psi_j(\cdot) \)'s and the \( \phi_j(\cdot) \)'s. The resulting partial demand functions are

\[
(22b) \quad \hat{h}_i^x(p,b,y_x) = y_x \frac{(\psi_i/p_i)^\sigma}{\sum \psi_j p_j^{1-\sigma}} + \frac{y_x^2}{f(p,b)} \left[ \frac{(\phi_i/p_i)^\sigma}{\sum \phi_j p_j^{1-\sigma}} - \frac{(\psi_i/p_i)^\sigma}{\sum \psi_j p_j^{1-\sigma}} \right], \quad i = 1, \ldots, N
\]

from which the share equations may be obtained by direct application of (12). Notice that these functions are no longer independent of \( y_x \) since the utility function is not homothetic.

The model actually estimated in Morey (1984) differs from (22b) because the variable \( x \) is substituted for the variable \( y_x \) - i.e. the partial demand functions and share equation take the form \( \hat{h}_i^x(p,b,x) \) and \( s_i(p,b,x) \). However, as pointed out earlier (p.9-6), this is an improper formulation. Yet, we can explain how something like this formulation, but not exactly the same, may arise. In forming the price variables, Morey includes both direct monetary costs \( (c_i, \text{say}) \) and the value of time, valued at the wage rate. Thus, \( p_i = c_i + wt_i \), where \( w \) is the wage rate and \( t_i \) is the time spent travelling and recreating at site \( i \). Since the partial demand functions and share equations are each homogeneous of degree zero in prices and income, one could normalize both prices and income by dividing them by \( w \), to obtain

\[
(23a) \quad \hat{h}_i^x(p,b,y_x) = \hat{h}_i^x \left( \frac{p_1}{w}, \ldots, \frac{p_N}{w}, b, \frac{y_x}{w} \right), \quad i = 1, \ldots, N
\]

9-14
(23b) \( s_i(p, b, y_x) = s_i\left(\frac{p_1}{w}, \ldots, \frac{p_N}{w}, b, \frac{y_x}{w}\right) \quad i = 1, \ldots, N. \)

Morey, in fact, does work with the prices \( (p_i/w) \) - i.e. prices measured in units of time rather than money. Moreover, he also measures recreation activity in units of time - \( x_i \) is one day of recreation at site \( i \), and \( x \) is the total number of days spent recreating over the season. But, from this one cannot legitimately infer that \( y \omega = x \) and, therefore one cannot substitute \( x \) as an argument on the right-hand sides of (23a, b). 8

Although the multinomial density attaches a non-zero probability to the event that \( t_i = 0 \) and can, therefore, be applied in practice to consumption data containing corner solutions, this does not necessarily make it a desirable tool for analyzing such data. In order for the density to be well-defined it is required that \( \pi_i > 0, i = 1, \ldots, N. \) If one identifies these parameters with the share system \( s_i(p, b, y_x), \ldots, s_N(p, b, y_x) \), as Morey (1981, footnote 14) does, rather than interpreting them as choice probabilities, this implies that the true demand for each good is positive and the only reason for observing zero shares in some particular data set is sampling variation rather than a structural feature of economic behavior - in the same way that when one tosses a fair coin several times it is possible through sampling variation to obtain a run of heads and no tails.
Admittedly, the fitted $s_i(\cdot)$ functions could take very small values, so that the expected consumption of any particular good is very small, but this is not a satisfactory solution to the problem of corner phenomena from an economic point of view. Assuming that there are corner solutions for reasons other than sampling variation, economic theory requires that $s_i(p,b,y_x) = 0$ for some range of $(p,b,y_x)$-space and, as we show in the next section, the internal structure of the $s_i(\cdot)$ functions changes when this occurs. The change in structure is not captured by (21b) or (22b), which are based on the presumption of an interior solution to the consumer's choice problem.

2. Share Models Used in the Literature

An alternative approach is to retain the multinomial model but interpret the parameters, $\pi_1, \ldots, \pi_N$, as choice probabilities arising from some structural economic model that explicitly incorporates the possibility of corner solutions. Thus, $\pi_i$ is
the probability that a budget share is positive and, even if the observed $s_i = 0$, one can have $\pi_i > 0$ and the way is cleared for an application of the density (18). The underlying economic model, however, is quite different from that which has been employed so far.

The economic model is different because, using the terminology introduced earlier, the macro-allocations now involve the determination of $x$, rather than $y_x$, and the micro-allocation involve the allocation of $x$ among individual sites. In addition the utility function is different in several respects, including the fact that it is no longer necessary to invoke the separability assumption that has been employed throughout this section. Thus, we work with the general utility function $u(x,b,z)$, and (although this is not crucial) we take $z$ as a scalar, representing the consumption of a Hicksian composite commodity whose price, $q$, is normalized to unity. We should emphasize that the assumption of an allocation of $x$, rather than $y_x$ is not compatible with the conventional economic framework of utility maximization (with two exceptions to be explained below) and should be interpreted as arising from a behavioral rather than an optimizing model of consumer behavior. The micro-allocations, however, are presumed to arise from maximizing behavior.

The micro-decision is to allocate a fixed total consumption, $x$, among $N$ different goods, where we recognize that the goods are indivisible and can only be consumed one unit at a time (e.g. one can only visit one recreation site at a time).
This indivisibility is explicitly recognized in the solution of the consumer's maximization problem, but in a special manner: instead of assuming that the consumer selects his entire portfolio of visits to the different recreation sites at a single instant (e.g. at the start of the recreation season), as is implied by conventional utility models, we now assume that he makes a separate choice of which site to visit each time he engages in the recreation activity. Given the predetermined total number of trips ("choice occasions"), \( x \), we introduce the choice vectors \( d_r = (d_{1r}, \ldots, d_{Nr}) \), \( r = 1, \ldots, x \), where \( d_{ir} = 1 \) if the \( i \)th site is selected on the \( r \)th trip and \( d_{ir} = 0 \) otherwise. Thus

\[
(24) \quad \sum_{i} d_{ir} = 1, \quad \sum_{r} d_{ir} = x_i, \text{ and } \sum_{r} d_{ir} = x.
\]

To allow for the possibility that the quality attributes or costs of visiting the sites vary over different choice situations, we can subscript these quality variables: \( b_r = (b_{1r}, \ldots, b_{Nr}) \),

where \( b_{ir} = (b_{1ir}, \ldots, b_{kir}) \), and \( p_r = (p_{1r}, \ldots, p_{Nr}) \). Also, let \( z_r \) denote the consumption of non-recreation goods and \( y_r \) the consumer's income on the \( r \)th choice occasion. Finally, let \( f(d_r, b_r, z_r) \) be the consumer's utility function relevant for the \( r \)th choice occasion. Given the macro-allocation decision, the micro-decision is to

\[
(25) \quad \text{maximize} \quad \sum_{r=1}^{x} f(d_r, b_r, z_r)
\]

subject to

\[
\sum_{j=1}^{N} p_{jr} d_{jr} + z = y \\
\text{r = 1, \ldots, x.}
\]

9-18
This maximization problem can be decomposed into x separate problems, that is, a separate decision problem for each choice occasion of the form

\[
\text{(26) } \maximize f(d_r, b_r, z_r) \text{ s.t. } \sum_{j} p_j d_j r^+ z_r = y.
\]

In order to describe the solution, suppose that on the \( r \)th choice occasion the individual has selected site \( i \). Conditional on this decision, his utility is

\[
(28) u_{ir} = f(0, \ldots, 0, 1, 0, \ldots, 0, b_{1r}, \ldots, b_{Nr}, y_r - p_{ir})
\]

\[
= v_i(b_{ir}, y_r - p_{ir}),
\]

if we assume that \( f(\cdot) \) satisfies weak complementarity. We will refer to \( v_1(\cdot), \ldots, v_N(\cdot) \) as conditional indirect utility functions. Since the consumer selects the site which yields the highest utility, the solution to (26) can be expressed in terms of these conditional indirect utility functions as

\[
(28) d_{ir} = \begin{cases} 
1 & \text{if } v_i(b_{ir}, y_r - p_{ir}) \geq v_j(b_{jr}, y_r - p_{jr}) \text{ all } j \\
0 & \text{otherwise}.
\end{cases}
\]

For estimation purposes, it is necessary to introduce a stochastic element into this demand model. In the context of discrete choices, such as arise here, this is commonly done by introducing a random element directly into the utility function producing what is known as a random utility maximization (RUM) model (See Hanemann, 1984a). The idea is that, although the consumer's utility function is deterministic for him, it contains some elements which are unobservable to the econometric investigator and are treated by the investigator as random.
variables. These elements will be denoted by the random vector $\mathbf{\varepsilon}$, and the utility function will be written $f(d_r,b_r,z_r;\varepsilon)$.

Under the RUM hypothesis, the consumer is assumed to maximize $f(d_r,b_r,z_r;\varepsilon)$ subject to the budget constraint in (26). If the RUM model satisfies the weak complementary condition, the conditional indirect utility functions have the general form $u_{ir} = v_i(b_{ir},y_{ir}-p_{ir};\varepsilon)$, $i = 1,\ldots,N$.

However, we assume that the random elements enter the utility functions in such a way that they, too, are affected by weak complementary, and we write the conditional indirect utility functions as a function of the scalar $\varepsilon_i$ rather than the vector $\varepsilon$:

$$ u_{ir} = v_i(b_{ir},y_{ir}-p_{ir};\varepsilon_i), \quad i = 1,\ldots,N. $$

(29)

The consumer's utility maximizing choice can still be expressed in terms of these conditional indirect utility functions along the lines of (28), except that the discrete choice indices $d_{1r},\ldots,d_{Nr}$ are now random variables with a mean $E[d_{ir}] = \pi_{ir}$ given by

$$ \pi_{ir} = \Pr\{v_i(b_{ir},y_{ir}-p_{ir};\varepsilon_i) \geq v_j(b_{jr},y_{jr}-p_{jr};\varepsilon_j) \text{ all } j\}. $$

(30)

In the applications by Hanemann (1978) and Caulkins et al. (1984), the random variables $\varepsilon_1,\ldots,\varepsilon_N$ are assumed to be independently and identically distributed extreme value variables, and the conditional indirect utility functions, (29), take the form
This generates the logit model of discrete choices

\[ u_{ir} = \bar{v}_i(b_{ir}, y_r - p_{ir}) + \varepsilon_i \quad i = 1, \ldots, N. \]

In the application to be presented in Chapter 10 we employ, instead, McFadden's (1978) Generalized Extreme Value Distribution

\[ p_r(\varepsilon_1 \leq s_1, \ldots, \varepsilon_N \leq s_N) = \exp\left[-G(e^{-s_1}, \ldots, e^{-s_N})\right] \]

where \( G \) is a positive, linear homogeneous function of \( N \) variables. When combined with (31) this yields discrete choice probabilities of the form

\[ \Pi_{ir} = e^{\bar{v}_i G_i(e^{-s_1}, \ldots, e^{-s_N})}/G(e^{-s_1}, \ldots, e^{-s_N}) \quad i = 1, \ldots, N \]

where \( G_i(\cdot) \) is the partial derivative of \( G(\cdot) \) with respect to its \( i \)th argument. In either case, the formulas for the choice probabilities may be substituted into the multinomial density (18) for maximum likelihood estimation of the parameters in the \( \bar{v}_i(\cdot) \) functions and any other parameters that may have been introduced into the joint density of the \( \varepsilon_i \)'s.

In estimating these parameters, the total number of trips, \( x \), is treated as an exogenous constant. However its determination is the subject of the macro-allocation decision, to which we now turn. We are aware of only two sets of circumstances in which the micro-decision allocating a fixed \( x \) among individual sites can be reconciled with conventional macro-allocations derivable from utility maximization. Suppose first...
that there is some unusual type of recreation activity in which one can participate only on a fixed number of occasions, $\bar{x}$, albeit at different sites - an example might be the temporary presence offshore of a rare species of fish or a heavily regulated fishery where the number of days of open season is a binding constraint on all users. People's preferences are such that they either do not participate in this recreation activity at all, or else participate $\bar{x}$ times (the maximum possible number of occasions). An individual may choose to visit different sites on different occasions, but for each person $\sum x_i = 0$ or $\sum x_i = \bar{x}$. Thus, in addition to the indivisibility in the "consumption" of individual sites, there is de facto an indivisibility in the overall level of participation in the activity. In these circumstances, a conventional utility maximization model could justify a micro-allocation decision of the type described above; the macro-allocation decision would be whether or not to participate in the activity, which is a discrete choice that could also be analyzed via a utility-theoretic logit or probit model. Indeed, both decisions could be derived from the maximization of a single underlying utility function based on (25). But the total number of visits to all sites if an individual did participate would still be an exogenous constant.

For most recreation activities this scenario is implausible since total site visitation, $x$, is a freely variable choice. In this case, the only way to fix $x$ in a manner consistent with an overall utility maximizing choice is to assume that $x$ emerges from a sequence of separate decisions: on each day of the
recreation season the individual decides both whether to participate in recreation on that day and which site to visit if he does participate. Suppose the season has $R$ days, and let $\theta_r$ represent the participation decision on the $r$th day, where $\theta_r = 1$ if the individual participates and $\theta_r = 0$ if he does not; as before, $d_{1r}, \ldots, d_{Nr}$ represents the choice of a site on the $r$th day, conditional on a decision to participate. In addition let $v_{Or} = v_0(y_r; \epsilon) \equiv f(0, \ldots, 0, b_{1r}, \ldots, b_{Nr}, y_r; \epsilon)$ measure the individual's utility if he does not participate in recreation on the $r$th day. His overall decision problem is to

\begin{equation}
\max_{\theta, d, z} \sum_{r=1}^{R} \left[ \theta f(d_r, b_r, z_r; \epsilon) + (1-\theta)v_0(y_r; \epsilon) \right]
\end{equation}

subject to

\begin{align}
& (36a) \quad d_r = 0 \text{ or } 1, \sum_i d_{ir} = 1, \quad r = 1, \ldots, R \\
& (36b) \quad \theta_r = 0 \text{ or } 1, \quad r = 1, \ldots, R \\
& (36c) \quad \sum_i p_{ir} d_{ir} + z_r = y_r, \quad r = 1, \ldots, R
\end{align}

which can be decomposed into $R$ separate problems of the form

\begin{equation}
\max_{\theta_r, d_r, z_r} \theta f(d_r, b_r, z_r; \epsilon) + (1-\theta)v_0(y_r; \epsilon)
\end{equation}

subject to the constraints in (36a,b,c). On any day the probability that the individual participates in recreation is given by

\begin{equation}
\pi_{Or} = P_r(\theta_r = 1) = P_r(v(y_r; \epsilon) \geq \max_i v_1(b_{1r}, y_r - p_{1r}; \epsilon), \ldots, v_N(b_{Nr}, y_r - p_{Nr}; \epsilon))
\end{equation}

while the probability that he visits site $i$, conditional on...
deciding to participate, is given by
\[ \pi_{i,r} = Pr\{d_{i,r} = 1|\theta_r = 1\} \]
\[ = Pr\{v_i(b_{i,r}, y_{r} - p_{i,r}; \epsilon) \geq v_j(b_{j,r}, y_{r} - p_{j,r}; \epsilon), j = 1, \ldots, N \} \]

The expected number of visits to all sites over the season is
\[ E[x_*] = \sum_{r=1}^{R} \pi_{i,r} \]

while the expected number of visits to the \( i \)-th site is
\[ E[x_{i,*}] = \sum_{r=1}^{R} \pi_{i,r} \pi_{i,R} \]

If \( \pi_{i,r} = \pi_i \) and \( \pi_{i,R} = \pi_o \) all \( r \), these become
\[ E[x_*] = \pi_o^R \]
\[ E[x_{i,*}] = \pi_i \pi_o^R. \]

The logic of this formulation is that the participation decision (which affects the macro-choice of how many trips to make over the season) and the site choice (the micro-decision) are interdependent and are made simultaneously by the individual. From the view of the econometrician they can be analyzed either simultaneously or, with some loss of efficiency but greater computational ease, separately. By analogy with (31) suppose that
\[ u_{o,r} = \bar{v}_o(y_r) + \epsilon_c \]

and let the joint density of \( \epsilon_0, \epsilon_1, \ldots, \epsilon_N \) be GEV with
\[ G(t_0, t_1, \ldots, t_N) = t_0 + [\Sigma (t_i^{1/(1-\sigma)})]^{1-\sigma} \]
where $\sigma \in [0,1]$ is, in effect, the common index of correlation for $\varepsilon_1, \ldots, \varepsilon_N$. Dropping the subscript $r$ for simplicity, the probability of selecting the $i$th site conditional on a decision to participate in recreation on any given day is

$$\Pi_i = \left( e^{\bar{v}_i/(1-\sigma)} \right) \cdot \left[ \sum e^{\bar{v}_j/(1-\sigma)} \right]^{-1} \quad i = 1, \ldots, N$$

(44) which parallels (32) except for the normalizing constant $(1-\sigma)^{-1}$. This constant cannot be identified from data on site choices alone; one can estimate only $\bar{v}_i/(1-\sigma)$, $i = 1, \ldots, N$. It is recovered, instead, from the data on the intensity of recreation participation. Define the "inclusive value", $I$, by

$$I \equiv \ln \left( \sum_{i=1}^{N} e^{\bar{v}_i/(1-\sigma)} \right) \quad (45)$$

The probability of participation in recreation on any day is

$$\pi_0 = e^{(1-\sigma)I}/(e^{\bar{v}_0} + e^{(1-\sigma)I}) \quad (46)$$

Given an estimate of $I$ from the analysis of site choices, the analysis of participation intensity based on (46) yields estimates of $\sigma$ and $\bar{v}_0$. This type of model has been applied to recreation demand by Feenberg and Mills (1980) and Caulkins et al. (1984), but with some differences. Substituting (46) into (40) and taking logarithms one obtains the regression model for total recreation activity

$$\ln x = \ln R - \ln(1 + e^{\bar{v}_0 - (1-\sigma)I}) \quad (47)$$

$$= \ln R - e^{\bar{v}_0 - (1-\sigma)I} \quad (48)$$

9-25
The estimation of this regression model by nonlinear least squares is an alternative to maximum likelihood estimation of a binary logit model based directly on (46). Feenberg and Mills (1980, p. 116) follow this route, but their regression model is somewhat different. For the purpose of analyzing site choices they set \( \sigma = 0 \) in (43) and (44) - i.e. they employ the standard logit model (32). Accordingly, their inclusive value index is

\[
(49) \quad I' = \ln(\sum_j \tilde{\nu}_j)
\]

and the above regression models would become

\[
(47') \quad \ln x_\ast = \ln R - \ln(1 + e^{\tilde{\nu}_0 - I'})
\]

\[
(48') \quad \approx \ln R - e^{\tilde{\nu}_0 - I'}.
\]

Instead, they estimate two alternative regression models of the form

\[
(50a) \quad x_\ast = -\tilde{\nu}_0 + \gamma I' \quad \gamma \neq 1
\]

\[
(50b) \quad \ln x_\ast = -\tilde{\nu}_0 + \gamma I' \quad \gamma \neq 1,
\]

neither of which is completely consistent with (47') or (48'). Caulkins et al. (1984) also set \( \sigma = 0 \) in their analysis of site choices, and they employ a binary logit model estimated by maximum likelihood for intensity of recreation participation. But, rather than originating from (46) with \( \sigma = 0 \), i.e.,

\[
(46') \quad \pi_0 = e^{I'}/(e^{\tilde{\nu}_0} + e^{I'}),
\]

9-26
their logit model is based on a participation probability of the form

\[ \pi_0 = \frac{e^I}{e^{\tilde{v}_0} + e^I} \]

where, instead of being given by (49), \( \tilde{I} \) is a linear function of the average price and quality characteristics of the various sites. Because of this difference, the site choices and the recreation participation decisions are not mutually consistent, in the sense of being derived from a single underlying utility maximization.

3. Alternative Model of Recreational Demand

As noted earlier, the above approaches to modelling total recreation activity assume that it emerges through a sequence of separate decisions—daily decisions made throughout the year, in the case of Caulkins et al.\(^18\). These decisions need not be independent, since the deterministic components of utility (i.e. the \( \tilde{v}_i(\cdot)'s \)) could be made to depend on previous choices during the recreation season, and the stochastic components (i.e. the \( \epsilon' \)s) might conceivably be correlated over time.

However, the decisions are uncoordinated in the sense that the individual never determines on any single occasion an overall allocation of time or money to recreation activity for the entire season. As a matter of modelling philosophy, we tend to find this unsatisfactory.

By analogy with \((15)\) in Chapter 8, we prefer to specify total recreation activity via some function of the form

\[ x_0 = H(\sigma(p), \tau(b), y) \]

or
where $\sigma(\cdot)$ and $\tau(\cdot)$ are some indices of the overall cost and quality of the recreation opportunities available to the individuals and $I'$ is the inclusive value index given by (49). We intend these as (somewhat arbitrary) behavioral relations, and recognize that they are not derivable from a hypothesis of overall utility maximization.

In estimating (52) or (53) we recognize that some individuals do not participate in recreation at all (i.e. $x_1 = 0$) and employ either Tobit analysis or Goldberger's (1964) two-stage approximation to Tobit. Consequently, the expected number of visits to all sites over the season may be cast in the form

\[(54) \quad E\{x_1\} = E\{x_1 | x_1 > 0\} \cdot \Pr\{x_1 > 0\},\]

where the second term on the right-hand size is the probability that the individual participates at all in recreation, and the expected number of visits to the $i$th site is given by

\[(55) \quad E\{x_i\} = \pi_i E\{x_1 | x_1 > 0\} \cdot \Pr\{x_1 > 0\}\]

where $\pi_i$ is given by (32) or (34).

Equations (54) and (55) may be contrasted with (40') and (41'). In particular it is important to note that the term $\Pr\{x_1 > 0\}$ in (54) and (55) is different from the probability $\pi_0$ defined in (38), (46) and (46'), which appears in (40') and (41'). The former measures the probability that the individual participates at all during the season, while the latter measures the probability that he participates on any given day. They are
related by

\[
Pr[x > 0] = 1 - (1 - \pi_0)^R.
\]

It is precisely because we believe that direct estimation of

\[Pr[x > 0]\]

is a more plausible method for dealing with

individuals who, because of old age, ill health, etc ne\text{ver} recreate that we prefer the modelling

approach based on the ad hoc macro-allocation functions (52) or

(53). If we wanted a fully utility-theoretic model of

participation and site choice, rather than (25) we would prefer

to employ the corner solution models to be described below.

Before proceeding to these models, it is important to point

out a practical implication of both the share models such as (20)

that were discussed in the earlier part of this section, and the

logit models such as (32) or (44) that we have just been

considering. This concerns the application of these models to

evaluate the welfare effects of events such as the closing of a

site or the improvement of a site's quality. The details of

these welfare calculations will be given in the final Section of

this Chapter. Here we wish to emphasize one aspect of them. The

logic of the share models, whether based on the Dirichlet,

logistic-normal, or multinomial distributions, is that one re-

covers the indirect (sub-)utility function \( \tilde{v}(p, b, y_x) \). Thus,

when we calculate compensating or equivalent variations from formulas

such as (16a,b) of Chapter 8 we obtain the compensation for the

entire season. By contrast, when we estimate logit models for

the allocation of visits among sites such as (32) or (44) and

perform welfare calculations as indicated below.
we obtain the compensation per choice occasion or per day of the recreation season. These compensations must therefore be multiplied by the number of choice occasions or the number of days in the recreation season in order to be comparable with the compensation estimates derived from the share models such as (20).

Corner Solution Models

In order to illuminate some of the problems which arise when one attempts to model corner phenomena in a manner fully consistent with utility theory, it is convenient to begin by describing how one models a special type of corner solution which Hanemann (1982a) has called an "extreme" corner solution. The utility maximization problem that concerns us in this section is:

\[
\text{maximize } u(x, b, z; \epsilon) \quad \text{s.t. } \sum p_c x_c + qz = y \\
x, z > 0.
\]

For simplicity we treat \( z \) as a scalar and set its price, \( q \), equal to unity. In contrast to the previous section, we are not now concerned with the possibility that commodities may be indivisible and we do not restrict the \( x_i \)'s to non-negative integer values. We are concerned, instead, with the non-negativity constraints in (57) and the circumstances in which they are binding. Extreme corner solutions arise when something...
in the structure of (57) forces a corner solution in which all but one of the \( x_i \)'s is zero - i.e. the consumer buys only one of the quality-differentiated goods. This can occur either because the utility function \( u(*) \) has a special structure which treats the \( x_i \)'s as perfect substitutes - examples are (21) and (22) of Chapter 8 - or because there is a set of additional constraints in (57) of the form
\[ x_i x_j = 0 \quad \text{all } i \neq j \]
i.e., for some logical or institutional reason the \( x_i \)'s are mutually exclusive in consumption. By contrast, a "general" corner solution arises when some, but not necessarily \( N-1 \), of the \( x_i \)'s are zero at the optimum. For most recreation choices one finds evidence of a general rather than an extreme corner solution; but, the analysis of extreme corner solutions will set the stage for more general models.

Since the modelling of extreme corner solutions has been discussed in some detail in Hanemann (1984a), we confine ourselves here to an outline of the main features. Suppose, for the moment, that the consumer has decided to consume only good \( i \) (visit site \( i \)). Invoking the assumption of weak complementarity, his utility conditional on this decision is
\[ u_i = u(0, \ldots, 0, x_i, 0, \ldots, 0, b, z; \epsilon) \equiv u^*_i(x_i, b_i, z; \epsilon). \]
Even given his selection of this site, he still has a decision to make - the number of times he should visit it over the recreation season. This decision is made by solving:
maximize $u^*(x_i, b_i, z; \varepsilon)$ s.t. $p_i x_i + z = y$

$x_i \geq 0, z \geq 0$.

The solution may involve setting $x_i = 0$ (i.e., he won't participate in any recreation over the season) or, less likely, $x_i = y/p_i$, he spends all of his income on recreation. Either of these corners may be handled by the methods to be described below, but for the moment we ignore them and simply write the solutions to (59) - the conditional ordinary demand functions - as $x_i = h_i^*(p_i, b_i, y; \varepsilon)$ and $z = z_i(p_i, b_i, y; \varepsilon) = y - p_i h_i^*(p_i, b_i, y; \varepsilon)$. The conditional indirect utility function obtained by substituting these functions back into $u^*_i(\cdot)$ is $v^*_i(p_i, b_i, y; \varepsilon)$.

Assuming that $u_i^*(\cdot)$ is a well-behaved direct utility function, these three functions possess all the standard properties. In particular $v_i^*(\cdot)$ is quasi-convex in $(p_i, y)$, decreasing in $p_i$ and increasing in $y$, and it satisfies Roy's identity

$$h_i^*(p_i, b_i, y; \varepsilon) = \frac{\partial v_i^*(p_i, b_i, y; \varepsilon)/\partial p_i}{\partial v_i^*(p_i, b_i, y; \varepsilon)/\partial y}.$$  

Under the RUM hypothesis, the quantities $x_i, z$ and $v_i^*$ are known numbers to the consumer but, because his preferences are incompletely observed, they are random variables from the point of view of the econometric investigator, and their distribution may be derived from the assumed joint density of $\varepsilon, f(\varepsilon)$.

All of the foregoing is conditional on the consumer's selecting site $i$. The discrete choice of which site to select (remember that only one site will be selected) can be represented by a set of binary valued indices $d_1, \ldots, d_N$ where $d_i = 1$ if $x_i > 0$ and $d_i = 0$ if $x_i = 0$. The choice may be
expressed in terms of the conditional indirect utility functions as

$$d_i(p, b, y; \varepsilon) = \begin{cases} 1 & \text{if } v_i^*(p_i, b_i, y; \varepsilon) \geq v_j^*(p_j, b_j, y; \varepsilon) \text{ all } j \\ 0 & \text{otherwise.} \end{cases}$$

For the observer, the discrete choice indices are random variables with a mean $E[d_i] = \pi_i$ given by

$$\pi_i(p, b, y; \varepsilon) = \Pr \{ v_i^*(p_i, b_i, y; \varepsilon) \geq v_j^*(p_j, b_j, y; \varepsilon), \text{ all } j \}$$

Now consider the original, unconditional utility maximization problem, which consists of (57) augmented, if necessary, by the constraints in (58). The unconditional ordinary demand functions associated with this problem will be denoted $h_i^i(p, b, y; \varepsilon)$, $i = 1, \ldots, N$ and $z(p, b, y; \varepsilon) \equiv y - \sum p_i h_i^i(p, b, y; \varepsilon)$, and the resulting unconditional indirect utility function is $v(p, b, y; \varepsilon)$. The relationship between these unconditional functions and the corresponding conditional ones is given by

$$h_i^i(p, b, y; \varepsilon) = d_i(p, b, y; \varepsilon) h_i^i(p_i, b_i, y; \varepsilon) \quad i = 1, \ldots, N$$

$$v(p, b, y; \varepsilon) = \max[v_1^*(p_1, b_1, y; \varepsilon), \ldots, v_N^*(p_N, b_N, y; \varepsilon)].$$
Given a set of data on observed consumption choices, the likelihood function can be constructed from (62) and (63) along the lines indicated in Hanemann (1984a). Here we wish to make three general points about this approach to modelling extreme corner solutions. First, the key building blocks are the conditional indirect utility functions, $v_1^*(\cdot), \ldots, v_N^*(\cdot)$. Once these have been specified, the discrete choice indices can be derived from them via (61), the conditional demand functions can be derived via (60), and the unconditional demand functions via (63). Thus we can construct an extreme corner solution model directly from the $v_i^*(\cdot)$'s without having to bother with the underlying direct utility function $u(x, b, z; \varepsilon)$. Second, the unconditional demand functions (63) embody an implicit switching regression model (i.e. a generalization of Tobit models), since they can be expressed equivalently in the form (using the case of $N = 2$ for simplicity):

$$
\begin{align*}
\text{if } v_1^*(p_1, b_1, y; \varepsilon) &\geq v_2^*(p_2, b_2, y; \varepsilon) \\
(x) &= \frac{\partial v_1^*(p_1, b_1, y; \varepsilon)}{\partial y} \\
&\quad - \frac{\partial v_2^*(p_2, b_2, y; \varepsilon)}{\partial y}, \\
\text{otherwise,}
\end{align*}
$$

In contrast the general (linear) single equation switching regression model takes the form

$$
\begin{align*}
\text{if } z \gamma + \eta \geq 0 \\
(65) \quad y &= W_1 \beta_1 + v_1 \\
&\quad + W_2 \beta_2 + v_2, \\
\text{otherwise},
\end{align*}
$$
where $Y$ is the dependent variable, $W_1$, $W_2$, and $Z$ are exogenous variables, $\beta_1$, $\beta_2$, and $\gamma$ are the coefficients to be estimated, and $\nu_1$, $\nu_2$, and $\nu$ are random error terms. The demand model (64) is clearly a special case of (65) where, because the discrete and continuous choices both flow from the same underlying utility maximization problem, the variables $W_1$ and $W_2$ are transformations of the variables $Z$, the coefficients $\beta_1$ and $\beta_2$ are directly related to the coefficients $\gamma$, and the random terms $\nu_1$ and $\nu_2$ are derived from the random term $\nu$. Thus, the random utility extreme corner solution demand model can be estimated by any of the statistical techniques developed for use with switching regression models while taking advantage of the additional restrictions inherent in the random utility formulation.

Our third point is a caveat: the practical application of these models rests on the ability to devise specific functional forms for the conditional indirect utility functions and the joint density $f_\epsilon(\epsilon)$ which yield reasonably tractable formulas for the discrete choice probabilities (56) and the conditional demand functions (60). This has in fact been accomplished: Hanemann (1984a) presents a variety of demand functions suitable for extreme corner solutions which offer considerable flexibility in modelling price, income, and quality elasticities. Several of these models are applied to the Boston recreation data set in Hanemann (1983a) for the subset of households - approximately one quarter of the sample - who visited only one site over the summer and, therefore, displayed evidence of an extreme corner solution in their behavior. The remaining households visited either no sites - which can also
be handled within the framework of an extreme corner solution model - or more than one site. However, none of the latter visited every site and, therefore, a general corner solution is required to model their behavior.

One approach to modelling general corner solutions is a straightforward generalization of that adopted above for extreme corner solutions. Instead of making the discrete choice the decision as to which site to visit, we can treat the decision to visit any combination of sites as a discrete choice. For example, suppose that the consumer decides to visit sites 2 and 3, but not sites 1 or 4, ..., N. Conditional on this discrete choice, his utility is \( u_{x_2}^2 = u(0, x_2, x_3, 0, ..., 0, b, z; \varepsilon) \)
\( \equiv u_{x_2}^3(x_2, x_3, b_2, b_3, z; \varepsilon) \), and he determines how many times to visit sites 2 and 3 by maximizing \( u_{x_2}^2(x_2) \) subject to \( p_2 x_2 + p_3 x_3 + z = y \), \( x_2 \geq 0 \), \( x_3 \geq 0 \), and \( z \geq 0 \). Denote the resulting conditional demand functions for these sites by \( h_{x_2}^2(p_2, p_3, b_2, b_3, y; \varepsilon) \) and \( h_{x_3}^2(p_2, p_3, b_2, b_3, y; \varepsilon) \) and the conditional indirect utility function by \( v_{x_2}^2(p_2, p_3, b_2, b_3, y; \varepsilon) \), and note that these functions satisfy Roy's Identity, (60). Proceeding similarly with all the other possible combinations of sites, the discrete choices can then be expressed in terms of the conditional indirect utility functions by a formula analogous to (61) and the unconditional demand functions are given by a formula analogous to (63). However, the problem with this approach is that, instead of \( N \) discrete choices, one now has to deal with \( 2^N \) discrete choices. In the recreation context, where \( N \) can easily be 20 or 30, this becomes extremely cumbersome. While the approach is formally correct in the sense that any other approach must yield equivalent results, one needs a simpler procedure.
An alternative procedure can be obtained by appealing to the economic considerations underlying the solution to the utility maximization problem (57), embodied in the Kuhn Tucker conditions. Substituting the budget constraint into the utility function, this problem may be written

\[(66) \quad \text{maximize } u(x, b, y - \sum_{i=1}^{N} x_i; \epsilon) \quad \text{s.t. } 0 \leq x_i \leq y/p_i, \quad i = 1, \ldots, N\]

and the Kuhn-Tucker conditions are

\[(67) \quad x_i > 0 \quad \text{as } \frac{\partial u(x, b, y - \sum_{j \neq i} p_j x_j; \epsilon)}{\partial x_i} - \frac{p_i \partial u(x, b, y - \sum_{j \neq i} p_j x_j; \epsilon)}{\partial z} > 0\]

Suppose one observed an individual who purchases quantities \(\bar{x}_1, \ldots, \bar{x}_Q\) of goods 1, ..., Q(\leq N), and \(y - \sum_{i=1}^{Q} p_j \bar{x}_j\) of the Hicksian composite commodity, but nothing of goods Q+1, ..., N, where \(\bar{x}_i > 0\) and \(y > \sum_{j=1}^{Q} p_j \bar{x}_j\). Define the N random variables \(\eta_1, \ldots, \eta_N\) by

\[(68) \quad \eta_i = \eta_i(x, p, b, y; \epsilon) = \frac{\partial u(\bar{x}, 0, b, y - \sum_{j \neq i} p_j \bar{x}_j)}{\partial x_i} - \frac{p_i \partial u(\bar{x}, 0, b, y - \sum_{j \neq i} p_j \bar{x}_j)}{\partial z}\]

\[i = 1, \ldots, N\]

and let \(f_\eta(\eta_1, \ldots, \eta_N)\) be their joint density, obtained from \(f_\epsilon(\epsilon)\) by an appropriate change of variables. By virtue of (67), the probability of observing this consumption event is given by

\[(69) \quad \Pr\left\{ \begin{array}{l} x_i = \bar{x}_i, \quad i = 1, \ldots, Q \\ x_i = 0, \quad i = Q+1, \ldots, N \end{array} \right\} = \Pr\left\{ \begin{array}{l} \eta_i = 0, \quad i = 1, \ldots, Q \\ \eta_i \leq 0, \quad i = Q+1, \ldots, N \end{array} \right\}\]

\[= \int_{-\infty}^{0} \cdots \int_{-\infty}^{0} f_\eta(0, \ldots, 0, \eta_{Q+1}, \ldots, \eta_N) d\eta_{Q+1} \cdots d\eta_N.\]

9-37
If the consumer purchased none of the goods, so that \( Q = 0 \) and \( z = y \), the probability of this event is

\[
(70) \quad \Pr(z = y) = \int_{-\infty}^{0} \ldots \int_{-\infty}^{0} f_\eta(n_1, \ldots, n_N) \, dn_1 \ldots dn_N
\]

while, if he purchased some quantity of every good (i.e. an interior solution), so that \( Q = N \) and \( y > \sum_{j=1}^{N} p_j x_j \), the probability is

\[
(71) \quad \Pr(x_i = \bar{x}_i, i = 1, \ldots, N) = f_\eta(0, \ldots, 0).
\]

Given an entire sample of consumers located at different corner solutions, the likelihood function would be the product of individual probability statements each having the form of (69), (70) or (71).

This approach to the modelling of general corner solutions was independently proposed by Hanemann (1978) and Wales and Woodland (1978). Two specific examples, both based on the Linear Expenditure System utility model, are

\[
(72) \quad u(x,b,z;\epsilon) = \sum_{j=1}^{N} \psi_j(b_j,\epsilon_j) \ln(x_j + \theta_j) + \ln z
\]

\[
(73) \quad u(x,b,z;\epsilon) = \sum_{j=1}^{N} \theta_j \ln[x_j + \alpha_j + \psi_j(b_j;\epsilon)] + \ln z
\]

where

\[
\psi_j(b_j,\epsilon_j) = \exp[\sum_{k} \gamma_{jk} b_{jk} + \epsilon_j] \geq 0 \quad j = 1, \ldots, N
\]
and the \( \theta_j \)'s, the \( \gamma_k \)'s and the \( a_i \)'s are the coefficients to be estimated, together with any parameters of the joint density \( f_v(\varepsilon) \).

Define the constants \( t_1, \ldots, t_N \) by

\[
t_i = -\sum_k \gamma_k b_{ik} + \ln \left( \frac{\theta_i (\bar{x}_i + a_i)}{\bar{y} - \sum_j p_j \bar{x}_j} \right) \quad i = 1, \ldots, Q
\]

(74)

\[
t_i = -\sum_k \gamma_k b_{ik} + \ln \left[ \frac{p_i \theta_i}{\bar{y} - \sum_j p_j \bar{x}_j} \right] \quad i = Q+1, \ldots, N.
\]

Then, in the case of the utility model (72), the probability statement (69) becomes

\[
\text{Pr}\left\{ x_i = \bar{x}_i \quad i=1, \ldots, Q \atop x_i = 0 \quad i=Q+1, \ldots, N \right\} = \int_{-\infty}^{t_{Q+1}} \cdots \int_{-\infty}^{t_N} f_v(t_1, \ldots, t_Q, \varepsilon_{Q+1}, \ldots, \varepsilon_N) d\varepsilon_{Q+1} \cdots d\varepsilon_N.
\]

In particular if the \( \varepsilon_i \)'s are independent extreme value variates, this probability has a closed-form expression:

\[
\exp \left( -\sum_i t_i \right) \prod_{i=1}^N \exp \left( -\sum_j \exp \left( -t_j \right) \right).
\]

In the case of the utility model (73), define \( t_1, \ldots, t_N \) by

\[
t_i = -\sum_k \gamma_k b_{ik} + \ln \left( \frac{\theta_i (\bar{x}_i - a_i)}{p_i + \sum_j p_j \bar{x}_j} \right) \quad i=1, \ldots, Q
\]

(77)

\[
t_i = -\sum_k \gamma_k b_{ik} + \ln \left[ \frac{\theta_i}{p_i + \sum_j p_j \bar{x}_j} \right] \quad i=Q+1, \ldots, N.
\]

The probability statement (69) becomes

\[
\text{Pr}\left\{ x_i = \bar{x}_i \quad i=1, \ldots, Q \atop x_i = 0 \quad i=Q+1, \ldots, N \right\} = \int_{t_{Q+1}}^{\infty} \cdots \int_{t_N}^{\infty} f_v(t_1, \ldots, t_Q, \varepsilon_{Q+1}, \ldots, \varepsilon_N) d\varepsilon_{Q+1} \cdots d\varepsilon_N.
\]

9-39
Once again, a closed form expression can readily be obtained when the $c_i$'s are independent extreme value variates. These utility functions and stochastic specifications by no means exhaust the possibilities, and we are currently actively exploring alternative formulations.\textsuperscript{19}

Two general points emerge from this analysis which are worth emphasizing. First, the probability expressions such as (78) generally require the evaluation of an $(N-Q)$-dimensional cumulative distribution function – i.e. a multiple integral whose dimensionality corresponds to one less than the number of commodities not consumed. In the recreation case, where $N = 20$ but a typical individual visits only $Q = 2$ or $3$ times, the evaluation of these integrals may be a daunting task, unless one invokes something like the independence assumption used above.\textsuperscript{20} The dimensionality $(N-Q)$ is fundamental, in that it is rooted in the logic of the utility maximization problem. Moreover, it represents an improvement over the approach, derived from the analysis of extreme corner solutions, which the discrete choices implied by the analog of (62) for general corner solutions involve, in principle, up to a $(2^N-1)$ dimensional cumulative distribution function.

Second, there is a basic tradeoff between achieving simplicity in the Kuhn-Tucker conditions.
and in the demand functions. This can be seen, for example, in the utility model (72). The marginal utility term appearing in the Kuhn Tucker conditions involves simply \( \psi_i(b_i, \varepsilon) \), whereas the \( i \)th demand function has the form

\[
(79) \quad x_i = -\theta_i + \left( \frac{\psi_i}{1 + \sum_j \psi_j} \right) \frac{1}{p_i} (y + \sum_j p_j \theta_j) \quad i = 1, \ldots, Q
\]

which involves the ratio \( \psi_i(1 + \sum \psi_j)^{-1} \). As we observed in the previous section, there is generally a choice between having a simple distribution for the random variables \( \psi_1, \ldots, \psi_N \) and having a simple distribution for ratios formed from them. In order to appreciate the significance of this tradeoff, it is necessary to consider the distinction between estimation and prediction as facets of the modelling activity. Both involve probability statements — estimation, for the purpose of forming likelihood functions; prediction, for the purpose of calculating the expected demand for sites under different price or quality regimes. In conventional demand analysis, including the share models described in the previous section, estimation and prediction are both based on essentially the same thing — the system of demand or share equations. Therefore, generally speaking, a stochastic specification which facilitates the process of estimation will also facilitate that of prediction, and conversely. As the above example illustrates, this is not true when we deal with corner solutions, where estimation can be based on the (perhaps simple) Kuhn-Tucker conditions, while prediction is based on the (perhaps complex) demand functions.
It would be convenient, therefore, if we could express the Kuhn-Tucker conditions not in terms of derivatives of the direct utility function, as in (67), but rather in terms of derivatives of the indirect utility function. Since Roy's Identity applies even in the face of corner solutions, there is some hope that a specification which achieves simplicity for the demand functions would also confer simplicity on the indirect utility Kuhn-Tucker conditions, if they existed. Furthermore, as discussed in the next section, welfare evaluation for price and quality changes are based on indirect utility functions, so that a specification which simplified their stochastic structure might also simplify the task of welfare evaluations. Is it possible, then, to analyze corner solutions exclusively in terms of indirect utility functions?

That it is, in fact, possible to do this has recently been proved by Lee and Pitt (1983) and Hanemann (1984d). Before describing the implications for the specification and estimation of random utility models, we shall summarize these theoretical developments, following the presentation in Hanemann (1984d). For simplicity we switch, temporarily, to a deterministic utility setting (i.e. we omit the vector ε) and we ignore the numeraire good, z, and the quality variables, b. Thus we can consider an indirect utility version of the Kuhn-Tucker conditions for the maximization problem:

\[
(80) \quad \max_{x} u(x_1, \ldots, x_N) \quad \text{subject to } \sum p_i x_i = y, \quad x_i \geq 0 \quad i = 1, \ldots, N
\]
As before, we denote the resulting ordinary demand functions by \( h^i(p,y) \), \( i = 1, \ldots, N \), and the indirect utility function by \( v(p,y) \). We shall now refer to these as the "constrained" demand and indirect utility functions, and we shall denote the vector of demands by \( \bar{x} \). In order to proceed, we need to introduce several additional maximization problems which are companions to (80). One of these problems is (80) minus the inequality constraints:

\[
(81) \quad \max_{\vec{x}} u(x_1, \ldots, x_N) \quad \text{subject to} \quad \sum p_i x_i = y.
\]

We refer to this as the "unconstrained" maximization problem, and we call the resulting demand functions, denoted \( h^i(p,y) \), and indirect utility function, denoted \( v^*(p,y) \), the "unconstrained" demand and indirect utility functions. As a shorthand, the consumption vector which solves (81) will be denoted \( x^* \). The logic of (81) is that the consumer is implicitly allowed to purchase negative quantities of goods, which is meaningless economically but serves as a useful artifact for our analysis.

The remaining utility maximization problems all involve equality constraints, of the form

\[
(82) \quad \max_{\vec{x}} u(x_1, \ldots, x_N) \quad \text{subject to} \quad \sum p_i x_i = y \quad \text{and} \quad x_i = 0
\]

or

\[
(83) \quad \max_{\vec{x}} u(x_1, \ldots, x_N) \quad \text{subject to} \quad \sum p_i x_i = y, \quad x_i = 0 \quad \text{and} \quad x_2 = 0, \text{ etc.}
\]

We refer to these as "partially constrained" problems and we denote the solution vectors by \( 1x^* \) or \( 12x^* \), etc., the demand functions by
$1^* h_1(\cdot)$ or $12^* h_1(\cdot)$, etc. and the indirect utility functions by $v_1^*(\cdot)$ or $12^* v^*(\cdot)$, etc. It is important to note that $1^* x_1 = 12^* x_1 = 12^* x_2 = 0$ (i.e. the first element of $1^* x$ and the first two elements of $12^* x$ are zero) and that not all of the prices appear as arguments in these partially constrained demand and indirect utility functions. Thus,

$$1^* x_i = 1^* h_1(p_2, \ldots, p_N, y), \quad i = 2, \ldots, N,$$

(84)

$$12^* x_i = 12^* h_1(p_3, \ldots, p_N, y), \quad i = 3, \ldots, N,$$

and

$$12^* u = 12^* v(p_3, \ldots, p_N, y).$$

With this notation in hand, we can now describe Hanemann's findings through a series of theorems whose proofs we omit. These involve the following regularity conditions on the direct utility function:

(D-1) $u$ is a continuous real-valued function defined over $\mathbb{R}^N$.

(D-2a) $u$ is non-decreasing in each argument, and is strictly increasing in at least one argument.

(D-2b) If $x_i = 0$ for some, but not all, indices $i$, $u$ is increasing in at least one argument $x_j$ where $x_j > 0$.

(D-3) $u$ is strictly quasiconcave.

The purpose of the theorems is to establish various relations between the solutions to the unconstrained or partially constrained problems.
Theorem 1. If \( x_i^* > 0 \) for \( i = 1, \ldots, N \), then \( \overline{x_i} = x_i^* > 0 \) for \( i = 1, \ldots, N \). Let \( u \) satisfy (D-1), (D-2a), and (D-3); then \( \overline{x_i} > 0 \) for \( i = 1, \ldots, N \) implies that \( x_i^* = \overline{x_i} > 0 \) for \( i = 1, \ldots, N \).

Theorem 2. Let \( u \) satisfy (D-1), (D-2a,b), and (D-3). Suppose \( \overline{x_i} = 0 \), \( \overline{x_i} > 0 \) for \( i = 2, \ldots, N \). Then (i) \( x_i^* \leq 0 \), and (ii) \( x_i^* = \overline{x_i} > 0 \) for \( i = 2, \ldots, N \).

Theorem 3. Let \( u \) satisfy (D-2a) and (D-3). Suppose that \( x_1^* \leq 0 \) and \( \overline{x_i} > 0 \) for \( i = 2, \ldots, N \). Then (i) \( \overline{x_1} = 0 \) and (ii) \( \overline{x_i} = \overline{x_i} > 0 \) for \( i = 2, \ldots, N \).

Given some set of indices, \( A \), let \( (A-i)x^* \) denote the solution to

\[
\max \limits_x u(x_1, \ldots, x_N), \quad \sum \limits_i x_i = y, \quad x_j = 0 \quad \text{all } j \in A, \quad j \neq i
\]

where it is understood that the index \( i \) is a member of \( A \).

Theorem 4. Let \( u \) satisfy (D-1), (D-2a,b), and (D-3). Suppose that, for some set of indices \( A \), \( \overline{x_i} = 0 \) for all \( i \in A \) and \( \overline{x_i} > 0 \) for all \( i \notin A \).

Then, (i) \( x_i^* \leq 0 \) for at least one index \( i \in A \)

(ii) \( (A-i)x_i^* \leq 0 \) for each \( i \notin A \)

(iii) \( A x_i^* = \overline{x_i} > 0 \) for all \( i \in A \).

Theorem 5. Let \( u \) satisfy (D-2a) and (D-3). Suppose that, for some set of indices \( A \), (a) \( x_i^* \leq 0 \) for at least one index \( i \in A \), (b) \( (A-1)x_i^* \leq 0 \) for each \( i \notin A \), and (c) \( A x_i^* > 0 \) for all \( i \notin A \).

Then, (i) \( \overline{x_i} = 0 \) all \( i \in A \)

(ii) \( \overline{x_i} = A x_i^* > 0 \) all \( i \notin A \).
Theorem 1 says that if, at some point in (p, y)-space, the unconstrained demand functions are all positive, \( h_i^*(p, y) > 0 \) \( i = 1, \ldots, N \), then the constrained maximization problem (80) has an interior solution, and conversely. Suppose, instead, that there is a corner solution. It turns out to make something of a difference whether the corner solution involves zero consumption of only one good or of several goods. If the former occurs, say only \( x_1 = 0 \), this implies that the unconstrained demand for good one is non-positive, while the partially constrained demands for all the other goods (i.e. the demands conditional on having \( x_1 = 0 \)) are all positive: \( h_1^*(p, y) \leq 0 \), \( h_i^*(p_2, \ldots, p_N, y) > 0 \) \( i = 2, \ldots, N \). This is the content of Theorem 2, while Theorem 3 states the converse. Similarly Theorems 4 and 5 cover the case where the corner solution involves zero consumption of several goods.

An important implication of part (ii) of Theorem 2 and part (iii) of Theorem 4 is that, at corner solutions, the demands for the goods which are being consumed are independent of the prices of those which are not: e.g. if \( x_1 = 0 \), the demand functions for the other goods, \( h_i(\cdot) \), \( i = 2, \ldots, N \), are independent of \( p_1 \) and depend only on \( p_2, \ldots, p_N \). In effect, when one takes the non-negativity constraints in (80) seriously, the constrained demand function is segmented, having the general form:

\[
(86) \quad h_1(p, y) = \begin{cases} 
    h_1^*(p_1, \ldots, p_N, y) & \text{if } (p, y) \in T_1 \\
    2h_1^*(p_1, p_3, \ldots, p_N, y) & \text{if } (p, y) \in T_2 \\
    23h_1^*(p_1, p_4, \ldots, p_N, y) & \text{if } (p, y) \in T_3 \\
    \vdots \\
    \text{etc.} 
\end{cases}
\]
where the events \((p,y) \in T_1, (p,y) \in T_2, (p,y) \in T_3\) etc. correspond to different sign patterns for the unconstrained and partially constrained demand functions, as specified in Theorems 1-5. (The exact rules for deriving \(U\) functions are given in Hanemann, 1984d). This is precisely the point that we made earlier in our discussion of share models: the corner phenomenon raises not only statistical issues (the piling up of a probability mass at \(x_i = 0\)), but also economic issues since it induces structural changes in the ordinary demand functions as indicated in (86).

In order to apply these theorems, however, we need to know how to obtain the various partially constrained demand functions. As Hanemann (1984d) shows, all that is required is a formula for the unconstrained indirect utility function, \(v^*(p_1, \ldots, p_N, y)\).

Writing this as a function of normalized prices, \(p_i^* = p_i/y\), the regularity conditions on \(v^*(p_1^*, \ldots, p_N^*)\) are that it be a continuous real-valued, quasi-convex function of \(N\) arguments and satisfy monotonicity conditions analogous to (D-2a,b). Everything else - both the constrained and partially constrained indirect utility functions and the constrained and partially constrained demand functions - can be derived from this function. The key is the observation that Roy's Identity applies to the unconstrained and all the partially constrained indirect utility functions, just as to the constrained indirect utility function. Thus

\[
x_i^* = -\frac{\partial v^*(p,y)/\partial p_i}{\partial v^*(p,y)/\partial y} \quad i = 1, \ldots, N
\]

\[
x_i^* = -\frac{\partial v^*(p_2, \ldots, p_N)/\partial p_i}{\partial v^*(p_2, \ldots, p_N)/\partial y} \quad i = 1, \ldots, N
\]

(87)
etc. This has two consequences. First, given $v^*$, $1v^*$, etc., the corresponding demand functions can be constructed via Roy's Identity in the usual manner. Second, since it can be shown that $\partial v^*/\partial y > 0$, $\partial_1 v^*/\partial y > 0$, etc., it follows that

$$\text{sign } (x_i^*) = -\text{sign}(\partial v^*/\partial p_i), \text{ sign}(1x_i^*) = -\text{sign}(\partial_1 v^*/\partial p_i), \text{ etc.}$$

Therefore, the conditions stated in Theorems 1-5 on the signs of $x_i^*$, $1x_i^*$, etc. can be translated into equivalent conditions on the signs of the price derivatives of the indirect utility functions $v^*$, $1v^*$, etc. For this reason, these theorems may be regarded as providing a set of indirect Kuhn-Tucker conditions, fully comparable to the conventional Kuhn-Tucker conditions, (67).

Finally, there are some simple rules for deriving the various partially constrained indirect utility functions from $v^*(p,y)$. For example, to obtain $1v^*(p_2,...,p_N,y)$ one solves

$$(88a) \quad \partial v^*(p_1,...,p_N,y)/\partial p_1 = 0$$

for $p_1 = \phi_1(p_2,...,p_N,y)$ and substitutes $\phi_1$ into $v^*$ to obtain

$$(88b) \quad 1v^*(p_2,...,p_N,y) = v^*[\phi_1(p_2,...,p_N,y),p_2,...,p_N,y].$$

Similarly, to obtain $12v^*(p_3,...,p_N,y)$ one solves

$$(88c) \quad \partial_1 v^*(p_2,...,p_N,y)/\partial p_2 = 0$$

for $p_2 = \phi_2(p_3,...,p_N,y)$ and substitutes $\phi_2$ into $1v^*$ to obtain
Proceeding in this manner we obtain the full set of \( 2^{N-1} \) partially constrained indirect utility functions. The constrained indirect utility functions can then be constructed in the same manner as the constrained ordinary demand functions, (86):

\[
12v^*(p_3, \ldots, p_N, y) = v^*[42(p_3, \ldots, p_N, y), p_3, \ldots, p_N, y]
\]

At this point we can switch back to the random utility framework employed earlier. Thus, we start with an unconstrained indirect utility function \( v^*(p_1, \ldots, p_N, b_1, \ldots, b_N, y; \varepsilon) \) and derive the various partially constrained indirect utility functions as shown in (88) noting that, if the underlying direct utility model is assumed to satisfy weak complementarity, the \( b_i \)'s will drop out along with the \( p_i \)'s as we pass from \( 1v^* \) to \( 12v^* \) to \( 123v^* \) etc. – For example, \( 1v^* \) takes the form \( 1v^*(p_2, \ldots, p_N, b_2, \ldots, b_N, y; \varepsilon) \). Next, we apply Roy's Identity to obtain the corresponding partially constrained demand functions. Finally we invoke Theorems 1-5 to obtain the requisite probability statements for the observed...
consumption outcomes. For example, if we observe an individual consuming positive quantities, \( \bar{x}_i \), of every good, where \( \bar{x}_i > 0 \), \( y > \sum_{i=1}^{N} p_i \bar{x}_i \), from Theorem 1 the probability of this event is

\[
\text{Pr}[x_i = \bar{x}_i, \text{all } i] = \text{Pr}[h_i^*(p, b, y; \epsilon) = \bar{x}_i, \text{all } i].
\]

Similarly, if we observe an individual consuming nothing of good 1 but positive quantities, \( \bar{x}_i \), of every other good, where \( \bar{x}_i > 0 \), \( i = 2, \ldots, N, y > \sum_{i=2}^{N} p_i \bar{x}_i \), from Theorems 2 and 3 the probability of this event is

\[
\text{Pr}\left\{\begin{array}{c}
x_1 = 0 \\
x_i = \bar{x}_i, \text{ for } i = 2, \ldots, N
\end{array}\right\. \right\}
\]

(91) \[
\text{Pr}\left\{\begin{array}{c}
h_i^*(p, b, y; \epsilon) \leq 0 \\
1 h_i^*(p_2, \ldots, p_N, b_2, \ldots, b_N, y; \epsilon) = \bar{x}_i \quad \text{for } i = 2, \ldots, N
\end{array}\right\}.
\]

When we come to the probability statements for more complex corner solutions there is an extra complication which has not so far been discussed. Theorems 4 and 5 state an equivalence between the event that \( x_i = 0 \) \( i = 1, \ldots, Q \), \( x_i = \bar{x}_i > 0 \), \( i = Q+1, \ldots, N \) and the following three sets of events:

(a) \( x_i^* \leq 0 \) at least one index \( i \in A = \{1, 2, \ldots, Q\} \)

(b) \( (A-i) x_i^* \leq 0 \) for each \( i \in A \)

(c) \( A x_i^* = \bar{x}_i \) \( i = Q+1, \ldots, N \).
essentially the same as the corresponding theorems in Lee and Pitt (1983), \(22\) the latter's versions of Theorem 4-5 omit conditions (a) and contain only conditions (b) and (c). Is condition (a) redundant? Hanemann proves that it is not. He shows that the following condition

\[(b') \quad (A-i)x_i^* \leq 0 \text{ for each } i \in A \text{ with at most one strict equality}
\]

implies (a) and would make it redundant; examples of this condition include

\[(b') \quad (A-i)x_i^* < 0 \text{ for each } i \in A\]

or

\[(b''') \quad (A-1)x_i^* = 0, \quad (A-1)x_i^* < 0 \quad i \in A, \; i \neq 1\]

But it is condition (b), not (b'), that is required in the proof of Theorems 4 and 5. In this respect, therefore, Hanemann's results differ from those of Lee and Pitt. However, in the context of a random utility model, since the two probabilities \(\Pr \{ (A-i)x_i^* \leq 0 \}\) and \(\Pr \{ (A-i)x_i^* < 0 \}\) are the same except on the space of measure zero, it follows that for all practical purposes we can dispense with condition (a) and write the probability of observing \(x_i = 0, \; i = 1, \ldots, Q\) and \(x_i = \overline{x}_i, \; i = Q+1, \ldots, N\), where \(1 < Q < N\), \(\overline{x}_i > 0\), and \(y > \sum_{i} p_i \overline{x}_j\), as

\[
\begin{align*}
\Pr & \{ \begin{array}{l}
x_i = 0 \quad i=1,\ldots,Q \\
x_i = \overline{x}_i \quad i=Q+1,\ldots,N
\end{array} \} \\
& = \Pr \{ (A-i)^{h_i^*}(p_{Q+1}, \ldots, p_N, b_{Q+1}, \ldots, b_N, y; \epsilon) \leq 0 \quad i=1,\ldots,Q \}
\end{align*}
\]
(92) are logically equivalent to those based on the direct Kuhn-Tucker conditions, such as (69), since they refer to the same event. Moreover, they are susceptible to the same problem of dimensionality since (92), like (69), requires in principle the evaluation of an $(N-Q)$ dimensional cumulative distribution function. However, there are cases where the indirect utility function $v^*(p,b,y;\varepsilon)$ does not have a closed-form representation as a direct utility function and therefore cannot be employed, whereas (92) is still available. There is also a direct link between this approach to the estimation of corner solutions and the approach that we mentioned earlier involving a discrete choice among $2^N$ alternatives (see page 36). It should be evident, for example, that the conditional demand and indirect utility functions introduced there correspond to the partially constrained demand and indirect utility functions associated with (82) and (83). However, the probability statements such as (92) provide a simpler approach to estimation than those based on the $2^N$ discrete choice probabilities.

Whatever the approach to estimation, we cannot escape the combinatorics implicit in the $2^N$ discrete choices when we come to construct the marginal probability distributions of the demands for individual sites, which would be needed to predict, say, the change in the expected demand for a site resulting from a change in its price or quality. Let $f_{x_1}(x_1)$ be the marginal density of $h_1(p,b,y,\varepsilon)$. Heuristically, from (86) this density will have the general form

$$(93) \quad f_{x_1}(x) = \int_0^{y/p_2} ... \int_0^{y/p_N} \Pr \left\{ x_1 = x, x_i = x_i^*, i = 2, ..., N \right\} d\bar{x}_2 ... d\bar{x}_N$$

where the probability statement inside the integral is given by expressions like (90), (91) and (92), depending on the region of
The evaluation of this marginal density and its mean, \( E[x_1] = \int x f_{X_1}(x) \, dx \), may require numerical techniques. But, as we mentioned earlier, there is an important distinction between complexity in the evaluation of an expression required for estimation and complexity in the evaluation of an expression required for prediction: because of the iterative nature of the estimation process, we can generally be far more tolerant of the latter than the former.

We are currently exploring various indirect utility models which may lend themselves to estimation on the basis of the indirect Kuhn-Tucker conditions, as an alternative to estimation based on the direct Kuhn-Tucker conditions. These include generalizations of the utility models employed in Hanemann (1983a, 1984a). However, dimensionality problems will always be a significant constraint, and in order to mitigate them we propose to employ utility models with a structure like that of (23) in Chapter 8. These models combine features of both extreme and general corner solutions since some goods (sites) are perfect substitutes for each other - therefore the consumer will select at most one of them - while other sites are not perfect substitutes and, therefore, may be consumed in any combination.

As a concrete example, suppose there are 15 sites which can be arranged into three groups: group A consists of sites 1, ..., 5; group B consists of sites 6, ..., 10; and group C consists of sites 11, ..., 15. We assume that the sites within each group are perfect substitutes, but that sites in different groups are imperfect substitutes. For this purpose, we employ the utility function (23)
of Chapter 8, although other formulations are certainly possible:

\[
(94) \quad u(x, b, z; \varepsilon) = \hat{u}\left[\sum_{i=1}^{6} x_i, \sum_{i=1}^{11} x_i, z + \sum_{i=1}^{15} \psi_i(b_i)x_i; \varepsilon\right]
\]

where \( \hat{u}(x_A, x_B, x_C, z; \varepsilon) \) is some existing four-argument utility function. Suppose we observe an individual who makes \( x_A \) visits to site 2, \( x_B \) visits to site 9, and \( x_C \) visits to site 11. The probability of observing this event is

\[
Pr\left\{ x_2 = \overline{x}_A, x_9 = \overline{x}_B \right\} = Pr\left\{ x_A = \overline{x}_A \right\} \text{ site 2 chosen out of group A } \nonumber
\]

\[
Pr\left\{ x_9 = \overline{x}_B \right\} = Pr\left\{ x_B = \overline{x}_B \right\} \text{ site 9 chosen out of group B } \nonumber
\]

\[
Pr\left\{ x_11 = \overline{x}_C \right\} = Pr\left\{ x_C = \overline{x}_C \right\} \text{ site 11 chosen out of group C } \nonumber
\]

\[
(95) \quad x \times Pr\left\{ \begin{array}{l}
\text{site 2 chosen out of group A} \\
\text{site 9 chosen out of group B} \\
\text{site 11 chosen out of group C}
\end{array} \right\}
\]

The second probability on the right-hand side of (95) takes the form of (62) and may have a relatively simple structure if we use some of the models in Hanemann (1984a). The first probability takes the form of (90), namely

\[
(96) \quad Pr\left\{ \begin{array}{l}
x_A = \overline{x}_A \text{ 2, 9 and 11} \\
x_B = \overline{x}_B \text{ selected out of}
\end{array} \right\} \nonumber
\]

\[
= Pr\left\{ \begin{array}{l}
h^*_A(p_2, p_9, p_{11}, b_2, b_9, b_{11}, y; \varepsilon) = \overline{x}_A \\
h^*_B(p_2, p_9, p_{11}, b_2, b_9, b_{11}, y; \varepsilon) = \overline{x}_B \\
h^*_C(p_2, p_9, p_{11}, b_2, b_9, b_{11}, y; \varepsilon) = \overline{x}_C
\end{array} \right\}
\]
where $h_A^*(\cdot)$, $h_B^*(\cdot)$ and $h_C^*(\cdot)$ are the unconstrained demand functions associated with $\hat{u}(x_A, x_B, x_C, z; \varepsilon)$. Similarly, if $x_C = 0$, i.e. the individual does not visit any site from group C, while $x_A > 0$ and $x_B > 0$, the probability of this event is

$$
\begin{align*}
\text{Pr}\left\{ x_2 = x_A, x_9 = x_B \right\} = \text{Pr}\left\{ x_A = x_A, \quad x_B = x_B \right\} \\
\text{Pr}\left\{ x_i = 0, \quad i \neq 2, 9 \right\} = \text{Pr}\left\{ x_C = 0 \right\}
\end{align*}
$$

The first probability on the right-hand side is given by

$$
\sum_{i=11}^{15} \text{Pr}\left\{ x_A = x_A, \quad x_B = x_B, \quad x_i = 0 \right\} = \text{Pr}\left\{ x_2 = x_A, \quad x_9 = x_B, \quad \text{site i chosen out of group C} \right\}.
$$

Each individual term in (98) having the following form (see 91):

$$
\begin{align*}
\text{Pr}\left\{ x_A = x_A, \quad x_B = x_B, \quad x_i = 0 \right\} = \text{Pr}\left\{ h_A^*(p_2, p_g, b_2, b_g, y; \varepsilon) = x_A \right\} \\
\text{Pr}\left\{ h_B^*(p_2, p_g, b_2, b_g, y; \varepsilon) = x_B \right\} \\
\text{Pr}\left\{ h_C^*(p_2, p_g, b_1, b_2, b_g, b_1, y; \varepsilon) \leq 0 \right\}.
\end{align*}
$$

Thus, the use of (94) reduces the effective dimensionality of the general corner solutions from $N=15$ to $N=3$, and hence to a discrete choice between $2^3$ rather than $2^{15}$ alternative combinations of consumption activities. We expect the gain in computational tractability to be very significant.
Welfare Evaluations

In this section we discuss some issues arising when the fitted multiple site demand models are used to derive money measures of the effect on an individual's welfare of a change in the prices or qualities of the available recreation sites, or of the closing of some site. We assume that the demand functions are compatible with the hypothesis of utility maximization—either at the micro-allocation level or at both the micro- and macro-allocation levels—so that the underlying indirect utility function can be recovered from them, and we are concerned with exact welfare measures rather than Marshallian approximations.

The basic theory of welfare measurement for quality changes was developed by Maler (1971, 1974) in the context of a deterministic (i.e. non random) utility function which ignored the possibility of corner-solutions. Given an indirect utility function, \( v(p, b, y) \), and some change in the set of prices and qualities facing an individual consumer from \((p', b')\) to \((p'', b'')\), two natural measures of the effect of this change on his welfare are the compensating and equivalent variations, \( C \) and \( E \), defined respectively by

\[
\begin{align*}
    v(p', b', y-C) &= v(p', b', y) \\
    v(p', b', y) &= v(p', b', y+E).
\end{align*}
\]

(100)

\( C \) and \( E \) measure not only the direction of the change in welfare, i.e.,

\[
\text{sign}(C) = \text{sign}(E) = \text{sign}[v(p'', b'', y) - v(p', b', y)].
\]
but also the magnitude of the change. The link between the $C$ and $E$
measures for pure quality changes and the conventional compensating
and equivalent variations for pure price changes is explored in
Hanemann (1980), where it is shown that standard results on the
sign of $(C-E)$ and the relation between $C$ or $E$ and the usual
Marshallian measure of consumer's surplus carry over from price to
quality changes. That paper also gives formulas, which are
summarized in Hanemann (1982a), for calculating $C$ and $E$ for a
quality change by setting up an equivalent price change and then
calculating the conventional compensating or equivalent variations
for this price change. This procedure can be applied when the
underlying utility function, $u(x,b,z)$, is obtained by the method of
transformations, as described in Chapter 7.

The task of performing welfare evaluations is more complex
when one works in a random utility setting. The theory of welfare
measurement in this context has been developed by Hanemann (1982c),
and revised and extended in Hanemann (1984c). We will provide a
sketch of this theory here, leaving the reader to refer to these
papers for a more detailed presentation. Both deal with extreme,
rather than general, corner solutions but these can involve either
purely discrete choices as in the logit models (32), (34), or mixed
discrete continuous choices. After summarizing the methodology for
these extreme corner solution models we will indicate how it can be
extended to cover general corner solution models of the type
discussed on pages 37-55.
The starting point for this welfare analysis, as for demand analysis, is the set of $N$ conditional indirect utility functions, $v_1(p_1, b_1, y; \varepsilon), \ldots, v_N(p_N, b_N, y; \varepsilon)$, from which the unconditional indirect utility function $v(p, b, y; \varepsilon)$ may be obtained via (64). This gives the utility attained by the individual maximizing consumer when confronted with the choice set $(p, b, y)$, which is a known number for him but a random variable for the econometric investigator. The cumulative distribution function $F_v(w) = \Pr(v(p, b, y; \varepsilon) < w)$ may be derived from the assumed distribution of the $\varepsilon$'s, $F_\varepsilon(\varepsilon)$, by a suitable change of variable. In particular, when the $v_i(\cdot)$'s have the form given in (31) - i.e. additive stochastic terms,

$$F_v(w) = F_\varepsilon(w - \bar{v}_1, \ldots, w - \bar{v}_N).$$

By analogy with (100), the compensations required by the individual to offset the change from $(p', b')$ to $(p'', b'')$ are given by

$$v(p'', b'', y-C; \varepsilon) = v(p', b', y; \varepsilon)$$

$$v(p'', b'', y; \varepsilon) = v(p', b', y+E; \varepsilon)$$

The problem in the random utility context is that $C$ and $E$ are now random variables, since they depend implicitly on $\varepsilon$. How, then, do we obtain a single number representing the compensating or equivalent variation for the price/quality change?

Hanemann (1984c) presents three different approaches to welfare evaluations in the random utility context, only one of
which has previously been recognized. That approach is based on the expectation of the individual's unconditional indirect utility function,

\[ V(p,b,y) = E[V(p,b,y;e)]. \]

In terms of this function, the measure of compensating variation is the quantity \( C' \) defined by\(^{24}\)

\[ (103) \quad V(p',b',y-C') = V(p',b',y) \]

This measure has been employed by Hanemann (1978, 1982c, 1983a), McFadden (1981), and Small and Rosen (1982). The formulas needed to calculate \( V(\cdot) \) for some common logit and probit additive-error random utility models are summarized in Hanemann (1982c). For example, in the GEV logit model (34),

\[ (104) \quad V(p,b,y) = \ln G(e^{\bar{V}_1}, ..., e^{\bar{V}_N}) + 0.57722... , \]

which is simply the inclusive value index (apart from Euler's constant, 0.57722...).

Another possible welfare measure is

\[ (105) \quad C^+ = E[C] \]

i.e. the mean of the individual's true (but random) compensation.

The distinction between \( C^+ \) and \( C' \) is subtle, but important. \( C^+ \)
is the observer's expectation of the maximum amount of money that the individual could pay after the change and still be as well off as he was before it. By contrast, \( C' \) is the maximum amount of money that the individual could pay after the change and still be as well off, in terms of the observer's expectation of his utility, as he was before it. A third possible welfare measure is derived as follows. One might want to know the amount of money such that the individual is just at the point of indifference between paying the money and securing the change or paying nothing and foregoing the change. For the observer, this could be taken as the quantity \( C^* \) such that

\[
(106) \quad P_r[\nu(p',b',y-C^*;c) \geq \nu(p',b',y;c)] = 0.5
\]

i.e. there is no more than a 50:50 chance that the individual would be willing to pay \( C^* \) for the change. It can readily be shown that, while \( C^+ \) is the mean of the distribution of the true compensation \( C \), \( C^* \) is the median of this distribution.

The procedures for calculating \( C^+ \) and \( C^* \) are described in Hanemann (1984c). Here we wish to emphasize that the three welfare measures, \( C' \), \( C^+ \) and \( C^* \), are in principle different, and the choice between them requires a value judgment on the part of the analyst. However, there are some circumstances in which some or all of them coincide. For example, in additive-error GEV models (which includes the standard logit model as a special case) Hanemann (1984c) proves that \( C' = C^* \). Similarly, in cases where the conditional indirect utility functions have the special form
where $Y > U$ is a constant that does not vary with $i$, he shows that $C' = C'$. Thus, when $\phi_i(\cdot)$ in (107) involves an additive stochastic term that is a GEV variate, $C^* = C' = C^*$. However, (107) is a highly restrictive assumption since it implies that both the discrete choices and the continuous choices (i.e., the conditional demand functions) are independent of the individual's income.25

If (a) there are income effects or (b) there are no income effects but the conditional indirect utility functions do not involve additive GEV variates, the difference between $C^+$ and $C^*$ can be substantial because the distribution of $C$, the true but random compensation, tends to be rather skewed, being bounded by zero at the low end but by income or $-\infty$ at the high end.

Thus its mean, $C^+$, may substantially exceed its median, $C^*$.

Hanemann (1983b, 1984c) also compares these welfare measures with an alternative calculation that was performed by Feenberg and Mills (1980) and Meta Systems (1983) using a logit model of recreation site choice. These authors were concerned with evaluating the benefits from an improvement in quality at an individual site—say, site 1. Thus, $b_1$ changes from $b_1'$ to $b_1''$ while $b_2, \ldots, b_N$ and $p_1, \ldots, p_N$ remain constant. Using the nonstochastic component of the conditional indirect utility function for this site, they calculated the quantity $\bar{C}$ defined by

\begin{equation}
\bar{C}(p_i, b_i, y; c) = \phi_i(p_i, b_i; c) + \gamma y \quad i = 1, \ldots, N
\end{equation}
By contrast, Hanemann shows that

\[(109) \quad C' = \Pi_1 \tilde{C} \]

and

\[(110) \quad C^+ = \Pi_1 \tilde{C} + \text{other terms.} \]

where \( \Pi_1 \) is the probability that the individual selects site 1 when faced with \((p_1, b_1, y)\).

The point is this: if we know for sure that the individual would select site 1, then \( \tilde{C} \) would indeed be the appropriate welfare measure. But, we can never be sure in the random utility context: there are only probabilities of site selection (i.e. \( \Pi_1 < 1 \), and we must weight the benefit \( \tilde{C} \) by the probability that the individual would have selected this site in the first place. In the data set used by Feenbery and Mills and Meta Systems, this probability was on the order of 0.05 to 0.2 for most sites. Equation (110) says that to the quantity \( \Pi_1 \tilde{C} \) we must add some other terms which measure the expected benefit to the individual if he originally selected some site other than 1 but switched to site 1 as a consequence of the improvement in its quality. In principle, the net effect of the two corrections to \( \tilde{C} \) implied by (110) is \( \tilde{C} \) could be larger or smaller than \( C^+ \). From some numerical simulations reported in Hanemann (1984c), it appears more likely that \( \tilde{C} \) overstates \( C^+ \) as well as \( C' \).
To this point our discussion has focused on welfare measurement in random utility models of extreme corner solutions. The theory of welfare measurement in general corner solution models is being developed, but the general outlines are clear. The distinction between the three welfare measures $C^+$, $C^*$, and $C'$ will carry over to general corner solutions, and it will still be true that $C^+ = C'$ only if there are no income effects. Because these random utility models typically involve something more complicated than the additive random structure, it is unclear whether the result that $C^* = C'$ can be obtained even when the random terms have a GEV distribution, as in the case of (77). However, the main difference between general and extreme corner solution models will be the greater complexity of the calculations, because the probability distribution of $v(p,b,y,e)$ - and hence of $C$ - that is implied by (89) is relatively cumbersome compared to (101).
FOOTNOTES TO CHAPTER 9

1. In part of this model - equation (3) - Sutherland treats each individual in his sample as the consumer; in the rest he treats the population of each "centroid" (a county, usually) as "the consumer". Furthermore he estimates separate demand systems for each of four recreation activities. Thus, there are four sets of $x_{it}$'s in the utility function, and four sets of demand functions corresponding to (3). These demand systems are independent, in the sense that the demand by consumer $t$ to visit site $i$ for recreation activity 1 is independent of his demand to visit the same site for recreation activity 2. Moreover, each of the demand systems is modelled in a similar manner. For simplicity, therefore, we focus on the demand system for a single activity and omit activity-specific subscripts.

2. One qualification should be noted: the partial demand system implied by a given share system is unique up to multiplication by an arbitrary positive constant. Thus, for any fixed $\lambda > 0$, $\lambda \tilde{h}_i(x)$ and $\tilde{h}_i(x)$, $i = 1, \ldots, N$, both imply the same share system.

3. Barten (1969) shows that the maximum likelihood estimates of $\Omega$ and the coefficients of $\tilde{h}_1^X(p, b, y_x), \ldots, \tilde{h}_N^X(p, b, y_x)$ are invariant with respect to which demand equation is treated as the residual.

4. Various assumptions may be made about the structure of $\Omega$: we could set $\Omega = \sigma^2 I_{N-1}$ (i.e. demands are independent and homo-
skedastic across sites), or $\Omega = \text{diag}(\sigma_1^2, \ldots, \sigma_N^2)$ (i.e. independent but heteroskedastic demands for sites).

Alternatively, we might assume that $\Omega$ is non-diagonal and employ Zellner's SUR estimation method. Finally, instead of assuming that $\Omega$ is constant across the individuals in the sample (i.e. different consumers' demands are independent), we might wish to parametrize $\Omega$ in some simple way so that it varies across the sample.

We allow for the dependence among $s_1, \ldots, s_N$ resulting from the constraint that $\sum_i s_i = 1$ by postulating that a subset $\{s_1, \ldots, s_{N-1}\}$ shares the $(N-1)$-dimensional multivariate normal, leaving the remaining share to be determined from the relation $s_N = 1 - s_1 - \cdots - s_{N-1}$.

5a. This stochastic specification implies that, for each consumer,

$$E[x_i] = \text{var}[x_i] = \bar{\pi}_i^X(p, b, y_x)$$

and $\text{cov}(x_i, x_j) = 0$, if $i \neq j$; hence the covariance matrix of the $x_i$'s, $\Omega$, varies across consumers. It also implies that $E[s_i] = s_i(p, b, y_x)$, and

$$\text{var}(s_i) = s_i(p, b, y_x)[1 - s_i(p, b, y_x)][1 + \sum_{j=1}^{N-1} \bar{\pi}_{ij}^X(p, b, y_x)]^{-1}$$

and

$$\text{cov}(s_i, s_j) = -s_i(p, b, y_x)s_j(p, b, y_x)[1 + \sum_{k=1}^{N-1} \bar{\pi}_{ik}^X(p, b, y_x)]^{-1}.$$}

6. With this stochastic specification, unlike the gamma/Dirichlet model, a consumer's demands for different sites need not be independent, but the covariance matrix is constant across individuals in the sample. Thus $E[\ln x_i] = \bar{\pi}_i^X(p, b, y_x)$,

$$\text{var}(\ln x_i) = w_{ij},$$

and $\text{cov}(\ln x_i, \ln x_j) = w_{ij}$, where $\Omega = \{w_{ij}\}$. Because of the off-diagonal terms in $\Omega$ compared to the covariance matrix for the gamma/Dirichlet model, the logistic normal distribution is likely to be more flexible for modelling shares, although, as Aitchison and Shen
(1983) show, any Dirichlet distribution can be closely approximated by a suitably parametrized logistic normal. The moments of the logistic normal exist and are finite but they do not have any simple form; however the moments of ratio $(s_i/s_j)$ or $\ln(s_i/s_j)$ can be expressed simply and are given by Aitchison and Shen. In the present context they take the form

$$ E[\ln(s_i/s_j)] = \frac{h_i^X(p,b,y)}{h_j^X(p,b,y)} $$

$$ \text{cov}[\ln(s_i/s_j), \ln(s_h/s_k)] = \sigma_{ih} + \sigma_{jk} - \sigma_{ik} - \sigma_{jh} $$

where $\sum \equiv (i,j)$. Assuming that $h_1^X(\cdot), \ldots, h_N^X(\cdot)$ satisfy (9), there is an additional complication with the lognormal/logistic normal model that does not arise with the gamma/Dirichlet model. For the latter, $E[p_i x_i] = \Sigma p_i E[x_i] = \Sigma p_i h_i^X(\cdot) = y_x$, whereas for the former

$$ E[\Sigma p_i x_i] = \Sigma p_i E[s_i] + \Sigma \Sigma p_i p_j \text{cov}[x_i x_j]. $$

Even if $\text{cov}[x_i, x_j] = 0$, all $i,j$, the right-hand side does not sum to $y_x$ because $E[x_i] = \exp[h_i^X(\cdot) + \frac{1}{2}w_{ii}]$. Thus the lognormal/logistic normal specification violates the summability condition on the demands, although it still satisfies the summability condition on the shares that $\Sigma s_i = 1$. This could be dealt with either by reparameterizing the normal distribution of the $x_i$'s or by dropping the assumption of weak separability and working with incomplete demand functions $h_i(p,b,q,y)$, which are not bound by (q).
Altonji (1986, pp. 156-7, p. 173) suggests some ad hoc devices for handling zero shares within the logistic normal distribution, but concludes that one may have to model them by introducing a probability mass at $x_i = 0$ and using the logistic normal as the conditional distribution of the non-zero shares, conditional on the others being zero. This is somewhat analogous to the approach employed in the utility-theoretic models of corner solutions to be discussed in the next section, where the probability mass at the boundary derives from economic considerations (i.e. the Kuhn-Tucker conditions holding as inequalities).

6b The Poisson distribution was considered by Hanemann (1978) for the Boston recreation data and rejected because the observed distributions of the $z_i$'s had both too long a right-hand tail and too much of a mass at zero.

7. It follows from (19a,b,c) that $E[x_i|x_\cdot] = \frac{h_i^x}{h}(p,b,y_x)$, $\text{var}(x_i|x_\cdot) = x_s(p,b,y_x)[1-s_i(p,b,y_x)]$, and $\text{cov}(x_i,x_j|x_\cdot) = -x_s(p,b,y_x)s_j(p,b,y_x) < 0$. This may be compared with the implications of the gamma/Dirichlet and lognormal/logistic normal models. Like the gamma, but not the lognormal models, the multinomial implies that the variance-covariance matrix of the $x_i$'s varies across the sample of individuals. However, it implies that, for any individual consumer, his demands (as opposed to shares) for different sites are negatively correlated, while the gamma implies that they are independent and the lognormal places no particular restriction on them. The negative covariance between $x_i$ and $x_j$ arises from the conditioning on the total $x_\cdot$, a feature which is absent from the gamma and lognormal models.

9-67
7a. One could obtain partial demand functions and share equations which were
equations of the form \( y_x \), rather than \( y_x \), only by assuming that, instead of the
typical utility maximization problem, the consumer solves

\[
\max_{x} \bar{u}(x, b) \text{ subject to } \sum x_i = x,
\]

where the resulting demand functions \( \bar{h}_i(b, x) \) and let \( \hat{h}_i(b, x) = \bar{h}_i / x \).
The relation between these functions and those employed earlier is

\[
\hat{h}_i(b, x) = \bar{h}_i(e, b, x,)
\]

\[
\hat{s}_i(b, x) = s_i(e, b, x).
\]

where \( e \) is a vector of \( N \) 1's.

8. A simple numerical example may help to illustrate the point.
Suppose there are two sites (\( N = 2 \)), with prices \( p_1 = $15 \)
and \( p_2 = $10 \), which include a time component valued at \( w = $8 \).
Suppose, also, that \( x_1 = 4 \) and \( x_2 = 2 \). Thus, \( x = 6 \)
(the individual spends six days on recreation), and \( y_x = $80 \)

\[
\frac{y_x}{w} = 10 \neq x.
\]

9. Indeed the CES utility model (21) precludes corner solutions since
it implies that all goods are essential; hence the
\( s_i(\cdot) \) functions in (21b) always satisfy \( s_i(\cdot) > 0 \). For
this reason the CES model seems unsuited to recreation
behavior: it is hard to believe that individual recreation sites
are all essential goods. The QES utility model (22a) does not
force any good to be essential and the demand function in (22b),
which is based on the presumption of an interior solution, can in
fact return negative or zero values for \( x_i \) and \( s_i \). As explained
in the next section, this is not the correct formula for the
demand function in its full generality. It should be noted that,
because the multinomial model presupposes that \( \rho > 0 \), the estimation
of (20) will in fact force the fitted \( s_i(\cdot) \) functions to be positive over
most or all sample data points.
10. Although we treat $z_r$ as the numeraire, the specification of these two variables for practical purposes is somewhat troublesome. Although it might seem natural to require that $\Sigma z_r = z$ and $\Sigma y_r = y$, it may not be desirable to assume that $z_r = z/x$ and $y_r = y/x$, because this implies also that, if $x$ changes, $y_r$ and $z_r$ change. The effective budget constraint facing the individual changes with $x$, which introduces an income-change component into welfare evaluations for price/quality changes, to be explained below. It would be more convenient to assume that $y_r$ is, say, monthly or weekly income (i.e. $y_r = y/12$, etc.) and correspondingly for $z_r$.

10a. We are implicitly assuming that the distribution $\varepsilon_t$ is different choice occasions. In principle we could subscript the random elements by the index $t$ and allow the distribution $\varepsilon_t$ to vary with $t$. However, in most applications we have no information on the temporal sequence of recreation choices over the season. Therefore, nothing would be gained by this generalization of the RUM model.

11. This simplification is not crucial; it is omitted, for example in random coefficients versions of the discrete choice model on the lines of Hausman and Wise (1978).
12. Morey (1981) also estimated this type of logit model in addition to his share model consisting of (20) and (21b). He compared each model's goodness of fit and concluded that the performance of the logit model was inferior. However, his logit model was based on conditional indirect utility functions of the form

$$V_i(b_{i1}, y-p_{i1}) = \beta (y-p_{i1}) + \sum_{k=1}^{4} \gamma_k b_{ik}$$

while the qualities indices $\psi_i(*)$ in his CES model, (21b), took the form

$$\psi_i(b_{i1}) = \sum_{k=1}^{4} \beta_k b_{ik}^{1/2} + \frac{1}{2} \sum_{k,h=1}^{4} \gamma_{kh} b_{ik}^{1/2} b_{ih}^{1/2}$$

where $\gamma_{kh} = \gamma_{hk}$. Thus, there were 5 parameters to be estimated in his logit model, but 15 parameters in his CES model (\sigma plus the 14 separate parameters in the $\psi_i$'s). The respective values of the maximized log-likelihood were -3528.001 and -3410.214. The comparison of the two models is complicated by the fact that they are not nested and, more importantly, by the difference in the number of parameters, which was not taken into account by Morey. Thus, we consider his conclusion to be somewhat premature.

13. One could also assume that the $\epsilon_1, \ldots, \epsilon_N$ have a multivariate normal distribution, which produces a probit model of discrete choice. However, the choice probability formulas corresponding to (33) involve an $(N-1)$ dimensional multivariate normal integral. In the application below where
In this application, as in Hanemann (1978) and Caulkins et al. (1984) it is assumed that \( p_{ir} = p_i \) and \( b_{ir} = b_i \), for all \( r \), but as demonstrated here that assumption is not essential.

14. In this context it would be natural to set \( y_r = y/R \) and \( z_r = z/R \). Note that we have assumed weak complementarity by making \( v_0(\cdot) \) independent of \( b_r \). However, \( v_0(\cdot) \) and, for that matter, \( f(\cdot) \) and \( v_1(\cdot) \) can contain characteristics of the individual such as age, sex, previous recreation experience, etc.

14a. One could impose further structure on \( \varepsilon_1, \ldots, \varepsilon_N \), allowing the stochastic terms associated with different groups \( \varepsilon_k \) to have different correlation indices.

15. It can be shown that \( I = E[\max[\bar{V}_1 + \varepsilon_1, \ldots, \bar{V}_N + \varepsilon_N]] \).

16. If \( \pi_0 > 0.5 \), it follows from (46) that \( \exp[\bar{V}_0 - (1-\sigma)I] < 1 \), which would justify the approximation \( \ln(1 + a) \approx a \).

Otherwise one could use some approximation such as \( \ln(1 + a) \approx a - a^2/2 \).

17. One can regard \( R \) or \( \ln R \) as being implicit in the intercept in (50a,b); the key differences between these equations and (47') or (48') lie in the functional form and the fact that \( \gamma \neq 1 \).

18. That is, their \( R \) appears to be 365. As pointed out in the preceding footnote the variable \( R \) does not appear explicitly in Feenberg and Mills' analysis, so one cannot tell whether they intend the recreation participation decisions to be made on a daily basis.

18a. This approach is adopted in Hanemann (1978).
19. For example, the approach adopted by Wales and Woodland (1978, 1983) is to assume that
\[ u(x,b,z;\epsilon) = \bar{u}(x,b,z) + \sum j \epsilon_j + z \]
where \( \bar{u}(\cdot) \) is a nonstochastic function. Hence the \( \eta_i(x,p,b,y;\epsilon) \) functions take the form

\[ \eta_i = \phi_i(x,p,b,y) + (\epsilon_i - p_i c_0) \quad i = 1, \ldots, N \]

where \( \phi_1, \ldots, \phi_N \) are nonstochastic functions. In this case it is convenient to assume that \( \epsilon_0, \ldots, \epsilon_N \) are multivariate normal. Observe that it simplifies the modelling task if one assumes that, unlike \( \partial u/\partial x_1, \ldots, \partial u/\partial x_N \), the term \( \partial u/\partial z \) does not contain a stochastic element, a simplification which was not exploited by Hanemann (1978) or Wales and Woodland.

20. If one assumed that \( f(\epsilon) \) were GEV, it might be possible to relax the independence assumption while retaining a closed form expression for (69).

21. The statistical feature, the piling up of a probability mass at \( x_i=0 \) is automatically incorporated in the utility-theoretic probability statements such as (70) or (92).

22. Hanemann offers a simplified proof of these theorems, employing a less restrictive set of regularity conditions on \( u \) than those of Lee and Pitt who require, in place of (D-2a,b), that \( u \) be strictly increasing in each argument. He also shows how the analogs of Theorems 1-5 apply to the dual problem of deriving the direct utility function from a given indirect utility function satisfying the regularity conditions mentioned earlier.
23. Thus, the conditional indirect utility function $v_{23}^c$ defined on page 36 is what we would write as $v_{i}^c$ in the notation of page 44.

24. At this point, the discussion will be conducted in terms of compensating variation. The analysis of equivalent variation measures is entirely analogous, and may be derived from that presented here by observing that the equivalent variation measure for a change from $(p', b')$ to $(p'', b'')$ is equal to the negative of the corresponding compensating variation measure for the change from $(p'', b'')$ to $(p', b')$.

25. This assumption of no income effects is employed by both McFadden (1981) and Small and Rosen (1982).


