NOTICE

AVIS

MEASURING CONSUMER SURPLUS FOR NONMONOTONIC CHANGES IN UTILITY

CHITRA RAJAN

submitted in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

at the

UNIVERSITY OF OTTAWA

December 1990

ISBN 0-315-70484-5

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

Dedicated to

Professor Thomas Brooks

in gratitude

CONTENTS

Part II

On the Measurement Problem

ACKNOWLEDGEMENTS

It has been an honor and a priveledge to have Professor S. Afriat supervise this work. I am grateful to him for this oppurtunity and for providing a most enjoyable learning experience. Members of the thesis committee made invaluable suggestions. Constructive criticisms from Professors Long and Bodkin are gratefully acknowledged. Professor Silva made thought provoking suggestions, which not only helped improve the thesis, but was a valuable education in itself. I am especially indebted to Professor Acheson for his encouragement and his sustained interest in my work.

I would also like to thank Dr. M. Hussain at the General Electric Research Center for all his help in the computational work, and for answering my numerous questions with such uncommon patience. I am grateful to Professor R. O'Neil, department of Mathematics at SUNY Albany, for introducing me to the grandeur of exterior calculus, and for allowing me the most generous use of his time.

My gratitude to all those who have had a profound impact on me over the years: Professor M. Demers who, through his own enthusiasm for the subject, inspired in me an interest in mathematics; my mother, my very first teacher; and my father, who taught me the joy of learning.

Above all, I am indebted to my husband, Krishna, who made this work possible, and to my children: Sonali, Anjana and Kishori, who made it all worthwhile.

ABSTRACT

We are concerned here with the problem of measuring consumer surplus from ordinary demand functions, in those situations where changes in the budget lead to variations in utility that are not monotonic. An example of such nonmonotonic utility changes are rectangular paths describing multiple price changes, where price increases are mixed with price decreases.

The literature on consumer surplus has primarily focussed on budget changes that lead to monotonic changes in utility, despite the common occurence of nonmonotonic paths in empirical work. We will show that many of the problems that arise for nonmonotonic paths either do not exist, or are insignificant when consumer surplus is measured for monotonic paths.

The first problem considered is the integrability question: the conditions which guarantee that the demand functions can be integrated to yield the generating utility function. We begin with a background chapter on the Slutsky matrix in a finite dimensional topological vector space. The link between the integrability conditions in terms of the Slutsky matrix and the Frobenius theorem in its classical form is derived in terms of differentiable manifolds. Alternatively, the problem is cast in terms of Grassmann manifolds, using the algebra of differential forms. A lemma of Poincare is used to obtain the conditions for path independent surplus measures.

The problem of obtaining a valid surplus measure that is consistent with consumer preferences is considered. It is shown that if the underlying preferences are nonhomothetic, the Marshallian surplus and the compensating variation are not valid utility indices, and may even give rise to incorrect ranking of the situations. Only the equivalent variation is a valid surplus measure for paths that lead to nonmonotonic changes in utility.

In most practical applications of welfare analysis, use is made of the Marshallian surplus, on the grounds that it varies only slightly along alternate paths, and along any particular path, will provide a good approximation to the Hicksian variations. While this is true for monotonic paths, we show that it is not so for paths that lead to nonmonotonic utility changes.

Such approximations are not necessary, as the equivalent variation can be calculated directly from ordinary demand functions. When preferences are quasihomothetic, expansion paths are linear, which allows us to calculate the integrating factor, and hence the indirect utility function. In the more general case, compensated changes in income are considered for the budget path, which gives us the expenditure function corresponding to the final utility level. This approach involves solving a differential equation. Closed form solutions as well as the method of numerical integration are considered.

As an example of paths that lead to nonmonotonic changes in utility, we consider the effects of a move to time differentiated electricity rates on the welfare of the consumer. We show that many of the studies use invalid methods to measure the changes in consumer surplus. Using data from time-of-day experiments in Connecticut, an exact measure of consumer welfare effects is calculated.

Introduction

The basic objective in the search for a surplus measure is to find one that will allow us to make welfare judgements about alternate equilibria, by telling us how an individual is affected when some variables affecting his decisions are altered. It is assumed that the changes are brought about by a policy, and the welfare judgement that we are looking for is the desirability of alternate policies.

We assume here that the consumer is the best judge of his own welfare, and that his choices in different situations are outcomes of a constrained utility maximization process. Hence it would be desirable to measure the surplus change in terms of his own actions, that is, in terms of his utility.

The problem however, is that the consumer's utility is not directly observable, and is in general, unknown. For example, when the consumer's budget changes, only the quantities consumed at these different prices and incomes are observed. Hence for purposes of applications, surplus measures have to be constructed from ordinary demand functions, in the process of which, numerous difficulties arise. We examine these problems in the context of changes in the budget of the consumer that lead to nonmonotonic changes in utility.

Since a policy change will be described entirely in terms of its effects on prices and incomes, much of the analysis will take place in the consumer's budget space, and certain topological properties of this space will be of particular interest.

Let E^n denote the n dimensional Euclidean coordinate space, consisting of n tuples of real numbers. Let x denote a commodity bundle, defined in the commodity space Ω^n,

$$\Omega^n = \{ \; x \in E_+^n; \; x \geq 0 \; \} \tag{1}$$

For every commodity in x, that is x_i, there exists a price p_i, defined in the price space Ω_n. Each element in Ω_n is a vector of n positive prices:

$$\Omega_n = \{ \; p \in E_{++}^n; \; p \gg 0 \; \} \tag{2}$$

The consumer also has a level of income, denoted by $m \geq 0$, defined in $E^1{}_+$.

The set of positive prices and income together form the *budget space*, denoted by Ω,

$$\Omega = \{ \; (p,m) \in E_+^{n+1}; \; p \in \Omega_n, \; m \in E_+^1 \; p \gg 0, \; m \geq 0 \; \} \tag{3}$$

We refer to the elements (p,m) as the budgets.

A *budget correspondence* can be defined as a mapping B from Ω to Ω^n. Its image is the budget set, and consists of elements which satisfy:

$$B(p,m) = \{ \; x : p^T x \leq m \; \} \tag{4}$$

1. *Defining a Policy Change*

Consider a policy, which if implemented, alters some of the variables facing the consumer. We assume here that the policy can be completely characterized in terms of its effects on the income of the consumer and the prices that he faces.

Let the policy lead to changes in the consumer's budget from an initial point (p^0,m^0) to a final point (p^1,m^1). Let $I = [a,b]$ denote this finite interval in E. We can then define a policy change formally as a mapping

$$(p^0,m^0) \in \Omega \mapsto (p^1,m^1) \in \Omega \tag{5}$$

from the initial vector (p^0,m^0) to the final vector (p^1,m^1) in the budget space Ω.

The direction of movement from the initial to the final point is crucial for much of our analysis. We follow Vartia's approach and view this movement as a sequence of continuous changes in prices and incomes, that can be partitioned sequentially:

$$a = t_0 < t_1 \ldots\ldots < t_k = b \tag{6}$$

(Vartia, 1983).

Let α be a differentiable curve connecting (p^0,m^0) to (p^1,m^1), that is, α: $[a,b] \mapsto E^{n+1}$. Let

$$\begin{aligned} \alpha^i(t) &= p_i(t) \qquad \text{for } i = 1 \ldots n \\ \alpha^{n+1}(t) &= m(t) \end{aligned} \tag{7}$$

be the parametric equations of α, for $t \in I = [a,b]$. A path is therefore a mapping

$$t \mapsto \alpha(t) \tag{8}$$

and defines the direction in which prices and income change in the budget space, as t moves in the interval I = [a,b].

For our analysis here, we require that the paths in Ω be continuously differentiable. To do so, we impose an additional topological property on the budget set B. We assume that B is an open polygonal set; and hence B is also *simply connected* (Apostol 1967). That is any path α' connecting the points (p^0,m^0) and (p',m') can be continuously deformed into another curve α'', also joining these two points, with all the intermediate curves during the deformation lying completely in B.

2. *Changes in Utility Due to Budget Changes*

The consumer is assumed to have a binary preference relation R on Ω^n, which is assumed to be connected and separable. If R is complete, reflexive, and transitive, then (from a theorem of Debreu), there e``ists a continuous, real valued utility function U,

$$U \; : \; x \in \Omega^n \mapsto \Re \tag{9}$$

which represents the preferences R such that

$$x\,R\,x' \Leftrightarrow U(x) \geq U(x') \tag{10}$$

Debreu, 1957. The utility function is a continous, quasiconcave function of class C^2, nondecreasing in x, and unique up to a strictly monotonic tranformation.

Under the hypothesis of rational behavior on the part of the consumer, maximization of utility, subject to the budget constraint, yields the demand function. As the consumer's budget changes, the quantities demanded, and hence the level of utility changes.

Although welfare comparisons can be made by computing the utility of alternate commodity bundles in Ω^n, it is often more convenient to compare in terms of his indirect utility associated with the various (p,m). We assume that the consumer has indirect preferences R^*, induced by the direct preference relation R:

$$(p,m)\ R^*\ (p',m') \quad \Leftrightarrow \quad x(p,m)\ R\ x(p',m') \tag{11}$$

The direct utility function U has associated with it an *indirect utility function* V, which is a mapping

$$V : \Omega \mapsto \Re \tag{12}$$

and is defined as

$$V(p,m) = \sup [\ U(x),\ x \in \Omega^n;\ p^T x = m\ ] \tag{13}$$

Since the solution to the constrained utility maximization is unique, the above equation can be written as

$$V(p,m) = Max\ [U(x);\ x \in \Omega^n;\ p^T x = m\ ] = U(x(p,m)) \tag{14}$$

(Afriat, 1980).

The indirect utility function is a continuous, finite function of class C^2. It is quasiconvex in p, and homogeneous of degree zero in (p,m). Given that U is nondecreasing in x, V is nonincreasing in p, and nondecreasing in m. Hence for all (p^0,m^0) and (p^1,m^1), if $V(p^0,m^0) > V(p^1,m^1)$, then the consumer's welfare is greater under (p^0,m^0) than under (p^1,m^1). Thus the change in consumer welfare can be measured in terms of the indirect utility function as well.

3. *Nonmonotonic Paths in the Budget Space*

Consider a sequence of changes in prices and income such that the changes in utility along this path is strictly increasing or decreasing. Such a path in the budget space along which preferences vary strictly monotonically has been defined by Zajac as a *complete strictly monotonic path* (Zajac 1979). We therefore define a nonmonotonic path in the budget space as one along which preferences both increase and decrease.

RECTANGULAR PATHS:

The term path refers to any continuous polygon curve of class C^1 defined for the interval $I = [a,b]$. Since B is assumed to be simply connected, any number of paths joining the initial and final points are possible. This places restrictions only on the initial and final points, but for any intermediate point $t \in I = [a,b]$, prices and income can take on any value. For example, if the price p_i does not change, it only implies that p_i^0 equals p_i^1; but the change in p_i, that is dp_i, does not have to equal zero for the entire path.

While the set of admissable paths include a large number of paths, most of them would be impractical in actual applications. For practical purposes, the typical choice is a rectangular path. These are step polygons where each side of the path is drawn parallel to the coordinate axes. Along such a path, prices and income move from their initial values, one at a time, while those that have moved stay at their new values (see Willig 1979). Hence these paths are easier to use in actual computations. It should be pointed out that for multiple price changes, several rectangular paths are possible, depending on the order in which prices are changed.

When a single price changes, or when several prices change but all the same direction, rectangular paths trace out paths in the budget space along which utility varies monotonically. Consider now the case of multiple price changes, where price increases are mixed with price decreases and rectangular paths are used to describe these changes. Utility would be increasing for segments of the path along which prices decrease, while decreasing for those segments of the path along which prices increase. Hence these are paths along which utility varies nonmonotonically.

4. Surplus Measures and the Line integral

The welfare change due to a policy is the difference in the consumer's utility between the initial and final points, measured in monetary units. Let Φ represent this money measure of surplus change. For the interval I = [a,b], defined by a path α, Φ is bounded by α, and is defined in the budget space Ω. The function Φ is of class C^1, and defines a vector field; that is, it assigns to each point in Ω a vector based at that point.

Consider a change in the consumer's budget from an initial point (p^0,m^0) to a final point (p',m'). The surplus change in money terms along a path c Φ is called a *line integral*

$$\int_c \Phi(\alpha(t),\ \alpha'(t))\ dt \tag{15}$$

defined as the integral of the vector valued function Φ over the interval [a,b]. The function Φ is an n+1 valued vector function, parametrically represented as

$$\Phi(\alpha(t)) = \Phi(\alpha^1,\ldots,\alpha^{n+1}) \equiv \Phi(p_1\ldots p_n,m) \tag{16}$$

The component functions of Φ, given by Φ^i, i = 1...n+1 are single valued functions defined at each point t on α, and form the coordinate functions of Φ. Hence the line integral can be written in its component form as:

$$\int_a^b \sum \Phi^i(\alpha(t)) \; \frac{d\alpha}{dt} \; dt \; = \; \sum \int_a^b \Phi^i(\alpha(t)) \; d\alpha^i \qquad \textbf{(17)}$$

5. The Marshallian Surplus

The early approach of Dupuit (1844) and Marshall (1890) was to measure the surplus change in terms of the changes in utility, converted into monetary units, using a 'converting factor' between money and 'utils'.

Consider the marginal utility of income λ:

$$\frac{\partial V(p,m)}{\partial m} \; = \; \lambda(p,m) \; > \; 0 \qquad \textbf{(18)}$$

Its reciprocal, that is $1/\lambda$, is the marginal cost of utility, and gives the number of dollars per 'util'. The Marshallian surplus uses this reciprocal to convert changes in utility into monetary units, and is given as $\Phi = 1/\lambda \; dV$, where Φ is a mapping

$$\Phi \; : \; \Omega_n \; x \; E^1 \; \mapsto \; \Omega \qquad \textbf{(19)}$$

The Marshallian surplus can be related to the ordinary demand function. Using Roy's identity and (18), we get:

$$\frac{\partial V(p,m)}{\partial p_i} \; = \; - \lambda \; x_i \qquad \textbf{(20)}$$

Hence the Marshallian surplus can be measured in terms of the ordinary demand functions plus any change in income between the two situations:

$$\int_c \Phi = \int_c \frac{1}{\lambda}\, dV = -\sum \int_a^b x_i(p,m)\, dp_i + m^1 - m^0 \qquad (21)$$

The problem with the Marshallian surplus is that the marginal utility of income is not observable and hence unknown. Unless λ is constant along the path with respect to the variables being changed, a unique correspondence between utility and money is not possible. It is well known that constancy of the marginal utility of income has implications that are quite restrictive (Samuelson, 1942). While the meaning of constancy of the marginal utility of income and its implications will be discussed in chapters I.2 and I.3, it should be pointed out that in the absence of these restrictions, the use of the Marshallian surplus is questionable.

5. The Hicksian Variations

Hicks' approach to the problem was to measure the value of a budget change to a consumer in terms of an equivalent change in income. This can be done using the expenditure function, e: Ω_n x $E^1 \mapsto E^1$

$$e(P,U^*) = inf\, [P^T X \ : \ U(X) \geq U^*] \qquad (22)$$

which is the minimum income that the consumer requires in order to achieve the level of utility U^*, given the prices p of the various commodities.

Introduction

The Hicksian variations can be measured as that amount of income, which at some reference price vector, affords the same level of utility as can be attained under a given budget constraint. This can be seen as a movement along an expansion path from the initial indifference surface to that corresponding to the final level of utility:

$$\Phi(\cdot) \;=\; de(p^{*},U) \;=\; e(p^{*},U^{1}) - e(p^{*},U^{0}) \tag{23}$$

Using the final price vector as the reference prices gives the compensating variation:

$$\Phi(C) \;=\; e(p^{1},U^{1}) \;-\; e(p^{1},U^{0}) \tag{24}$$

Similarly the equivalent variation is obtained using the initial price vector as the reference prices:

$$\Phi(E) \;=\; e(p^{0},U^{1}) \;-\; e(p^{0},U^{0}) \tag{25}$$

Since the Hicksian variations measure the surplus change in terms of the value of money before and after the budget change, these measures do not require that the marginal utility of income remain constant.

The Hicksian variations can also be defined as the amount of expenditure required to be on the same indifference surface for the change in prices from the initial to the final levels. This relates the Hicksian variations to areas left of the Hicksian demand functions. The latter is a mapping

$$h \;:\; \Omega_n \, x \, \Re \;\mapsto\; \Omega^n \tag{26}$$

defined so that any changes in prices has compensating changes in income to keep the consumer at some fixed level of utility. The Hicksian variations can be written as

$$\Phi(\cdot) \;=\; (m^1 - m^0) \;-\; \sum \int h_i(p, U^*) \; dp_i \tag{27}$$

where U^* equals U^0 for the compensating variation and equals U^1 for the equivalent variation.

Given the different surplus measures, we are concerned with the problem of finding an appropriate one for budget changes that lead to nonmonotonic changes in utility. Further, the surplus measure has to be constructed purely from information that is observable, that is the ordinary demand functions.

Part I

SOME THEORETICAL ISSUES

I.1

The Integrability Problem

Given that a consumer's preference ordering can be represented by a utility function, we can analyze the welfare of a consumer in different situations in terms of utility, measured in monetary units. However, the preferences of the consumer are not directly observable, and are unknown. Surplus measures therefore have to be constructed from information that is available; that is, the ordinary demand functions.

Here the theoretical problems that arise in constructing surplus measures from demand functions are examined. The first issue that we address is the integrability problem. Since the surplus measure is constructed from the gradient of the utility function, we need to find conditions sufficient to assure that the demand functions can be integrated to yield the generating utility function.

The problem of integrability is a familiar topic in utility theory, and its long history, which goes back to Antonelli in 1886, has been well documented (Hurwicz, 1974; Hurwicz and Uzawa, 1974; Afriat, 1980). Central to the integrability question is the Slutsky matrix. The conditions for the integrability of the direct demand functions in the budget space require that the Slutsky matrix be symmetric and negative definite. This allows the demand functions to be integrated to yield an integral surface, which is the income compensation function.

As a background to the integrability problem, we begin with a chapter on the Slutsky equation, which states that the observed variations in the quantities demanded due to a price change can be decomposed into a substitution and income effect. A unique approach to the description of this decomposition can be found in Afriat 1980, who has shown that the Slutsky matrix by being the product of a Jacobian and a *projector,* translates a decomposition in the budget space, into a corresponding decomposition in the commodity space.

In chapter 1.1, we prove that the Slutsky matrix is the product of a projector and a Jacobian. To begin with, a finite dimensional space is considered, whose elements are the inner product of the commodity and the normalized price vectors. We prove that this topological vector space can be uniquely written as the sum of two complementary linear subspaces. A *projector,* or a linear transformation of this topological vector space to its subspace is then considered, and it is shown that the matrix of this projection is a square one, which is idempotent. The link between projectors and idempotent matrices is established from certain standard properties of vector spaces, from which the results of Afriat follow.

In chapter 1.2 we address the integrability problem in its classical form. The sufficient conditions for the integrability of demand functions are obtained from a standard theorem of differential equations, as related to Frobenius (Spivak, 1970). The link between the Frobenius theorem in its classical form and the condition of Slutsky symmetry is explained by approaching the integrability question as an application of a problem in differential geometry.

In chapter 1.3, an alternative approach to deriving the integrability conditions is considered. This approach uses *Grassmann algebra* to define a *differential form* on a manifold. (Grassmann algebra has been used in the context of general equilibrium by Balasko (1988)). Here the special properties of this algebra is used to characterize the integrability of a distribution, and this constitutes the dual form of the Frobenius theorem.

1.1 The Slutsky Matrix, Idempotent Matrices, and Projectors

The Slutsky equation, which states that the observed variations in quantities demanded due to a price change can be decomposed into a substitution and income effect, is well known in consumer theory. Afriat (1980) has shown that the decomposition of a budget differential in the budget space into the sum of a substitution and an expansion term can be translated into a corresponding decomposition in the commodity space, the key to this translation being the Slutsky matrix.

The purpose here is to prove that the Slutsky matrix is the product of a Jacobian and a projector, and hence permits such a translation to take place. The proof lies in showing that the matrix of a linear transformation of a topological vector space to its subspace results in a projector. We do this by showing that if a topological vector space can be written uniquely as the sum of two complementary subspaces, then the matrix of such a projection is idempotent. The link between idempotent matrices and projectors is then established through certain properties of linear spaces and matrices.

The analysis is conducted in a normalized budget space. Given that $m > 0$, we may divide all prices by m and define the vector of normalized prices as $g = m^{-1}p$. For the sake of notational convenience, let us define g as a row vector, that is $g = m^{-1}p^T$. This space is a replica of the price space, and hence can also be represented by Ω_n.

From the Marshallian demand function x = x(p,m), defined as

$$x \ : \ (p,m) \in \Omega \ \mapsto \ \Omega^n \in E^n \tag{1}$$

is derived a function

$$f(g) \ = \ x(g,1) \tag{2}$$

which we define as the 'normal' demand function, whose values are $x \in \Omega^n$, with arguments $g \in \Omega_n$, and f is a mapping

$$f \ : \ g \in \Omega_n \ \mapsto \ \Omega^n \in E^n \tag{3}$$

The spaces Ω^n, Ω_n are dual and symmetric; that is, each represents the space of positive homogeneous linear functions defined in the other.

The normal demand functions, like the Marshallian demand functions, are mappings of class C^1, and are differentiable at every point in Ω_n. Let $L(\Omega_n, \Omega^n)$ denote the set of linear mappings from Ω_n to Ω^n. The differential of f at some point $q \in \Omega_n$, denoted by Df(q), is a linear mapping of Ω_n into Ω^n,

$$Df \ : \ \Omega_n \ \mapsto \ L(\Omega_n, \Omega^n) \tag{4}$$

Since g is itself a differentiable function of its arguments, we can consider the composition of differentiable mappings. Every component x_i of f(g) has an associated set of n mappings

$$x_{ip} \ = \ x_i(g(p,m)) \ \cdot \ g_p^T(p,m) \tag{5}$$

which gives the elements of the matrix

$$X_p \ = \ X_g \cdot g_p^T \tag{6}$$

where X_g is an n x n Jacobian matrix and

$$g_p^T = \frac{\partial g_i^T}{\partial p_j} = \delta_{ij} m^{-1} \tag{7}$$

δ_{ij} being the Kronecker delta, $\delta_{ij} = 0$ for $i \neq j$, and $\delta_{ij} = 1$ for $i = j$. Hence $g^T{}_m = 1\ m^{-1}$, and we have

$$m\, g_p^T = 1 \tag{8}$$

and

$$X_p\, m = X_g \tag{9}$$

Similarly, a mapping

$$x_m = X_g \cdot g_m^T \tag{10}$$

is obtained, where the components of the vector $g_m{}^T$ are

$$\begin{aligned} \frac{\partial g_i^T}{\partial m} &= \frac{\partial (m^{-1} p_i)}{\partial m} \\ &= -\, m^{-2} p_i^T \\ &= -\, m^{-1} g_i^T \end{aligned} \tag{11}$$

Hence

$$\begin{aligned} m\, g_m &= -\, g \\ x_m\, m &= -\, X_g\, g^T \end{aligned} \tag{12}$$

The path described by x, as m varies while p remains fixed is the *expansion paths* for prices p.

Consider a change in income from an initial level m by an amount v. We define the term $\tau = 1+v$ as the expansion factor or the amount of expansion.

Any price vector, with constant price ratios, lies on a line in the price space, that is, a ray through p. Since x is homogenous in (p,m), the changes in x due to changes in m can be seen to move along this ray through p.

The expansion path for the Marshallian demand function, is the linear operator which maps Ω into $L(\Omega, \Omega^n)$,

$$\begin{aligned} \dot{x} &= \frac{d}{d\tau} x(m,\ (\tau-1)m,\ p)\,|_{\tau=1} \\ &= x_m m \end{aligned} \tag{13}$$

which represents the rate of change of x as one departs from m in the direction of p.

Similarly, the expansion derivative of g

$$\dot{g} = \frac{d}{d\tau} g(m,\ (\tau-1)m,\ p)\,|_{\tau=1} \tag{14}$$

which, from (12), gives

$$\dot{g} = (-g\,m^{-1})\,m = -g \tag{15}$$

Hence for a normal demand function $x = f(g)$, the elements $\partial(f_i(\tau-1)\ g)/\partial\tau|_{\tau=1}$, define the vector

$$\begin{aligned} \dot{x} &= \frac{d}{d\tau} f(m,\ (\tau-1)m,\ g)\,|_{\tau=1} \\ &= -x_g\, g^T \end{aligned} \tag{16}$$

which is the expansion path for the normal demand function, and is identical with the vector in (13).

Consider a space W in E^n, with elements $k = gx$. Since for any two elements $k^0 = (gx)^\circ$ and $k^1 = (gx)'$ in W, and a scalar $\theta \in \mathfrak{R}$, the mappings $(k^0,k^1) \mapsto k^0+k^1$ and $(\zeta,a) \mapsto \theta a$ are continuous, the space W is equipped with a product topology that is compatible with the vector space structure. Hence W is a finite n dimensional topological vector space under its natural topology.

Let W_1 and W_2 be two linear subspaces of W, of dimensions n and m respectively, such that W is the algebraic direct sum of W_1 and W_2, denoted by $W = W_1 \oplus W_2$. Then the two spaces W_1 and W_2 are said to be *complementary*. For any subspace W_1, a complementary subspace can always be found: choose a basis (say G) for W_1, expand it to a basis of W, and let W_2 be the linear span of the basis vectors not in W_1. If H denotes the basis of W_2, then it can be seen that the join of the basis (G,H) is also a base, and the span of (G,H) is the union of W_1 and W_2, which is equal to the entire space W.

Given that a space W can be written as the direct sum of two complementary subspaces W_1 and W_2, we now wish to represent each vector k = gx in W uniquely as y + z, for $y \in W_1$ and $z \in W_2$. However, unless $W_1 = W$ or $W_1 = \{0\}$, an infinite number of subspaces exist which are complementary to W_1. Hence we have to ensure that the vectors y and z do belong to the vector spaces W_1 and W_2.

Consider the space W of dimension n over the field of real numbers. Any set of q vectors of gx in W is taken to compose a matrix Q of order n x q, which constitutes a vector array in W, with the vectors k = gx as elements. The span of these vectors is the subspace W_1 of W. Let us denote the vectors in W_1 by y = Qy, having the q vectors as generators.

For the subspace W_1, let s be the array of the maximal set of the independent vectors of W_1, any such set having n members Then s defines a base for W_1, such that the vectors y of W_1 are uniquely given by $y = sy$.

We invoke the following theorem by Afriat (1956):

THEOREM 1: A necessary and sufficient condition for a vector y to belong to a subspace W_1, with base s is

$$\begin{aligned} I_s y &= y \\ \text{where} \quad I_s &= s(s^T s)^{-1} s^T \end{aligned} \tag{17}$$

Then we obtain:

$$\begin{aligned} y &= sy \\ \text{where} \quad y &= (s^T s)^{-1} s^T y \end{aligned} \tag{18}$$

A necessary and sufficient condition that y belongs to the orthogonal complement of W_1 with base s is

$$\begin{aligned} I_s y &= 0 \\ \text{or} \quad J_s y &= y, \quad \text{where } J = 1 - I_s \end{aligned} \tag{19}$$

Moreover, the dimension of W_1 is equal to the rank of I_s, which is equal to the trace of I_s.

Consider the space W, with a subspace W_1. Fix another subspace W_2, such that $W = W_1 \oplus W_2$; $W_1 \cap W_2 = \{0\}$. If s_1 and s_2 are the bases of W_1 and W_2 respectively, then the join of their bases defines a regular square matrix $A = (s_1, s_2)$ which forms the base for the vector space W.

Since the space W is a topological vector space, equip W_1 and W_2 with the relative topologies. Then the space $W_1 \times W_2$ is equipped with the product vector space structure. Consider a mapping:

$$T : W_1 \times W_2 \mapsto W \tag{20}$$

defined by $T(y, z) = y + z$. Evidently T is linear, surjective, (because $W_1 \oplus W_2 = W$), and injective (because $W_1 \cap W_2 = \{0\}$), that is T is a vector space isomorphism. Thus the space W is structurally decomposable, with W_1 and W_2 as factors. Further, since the mapping $T : W_1 \times W_2 \mapsto W$ is bicontinuous, W is the topological direct sum of W_1 and W_2, and hence W_1 and W_2 are topological complements. Hence if v, y, and z are the elements of W, W_1, W_2 respectively, then any vector v in W has a unique resolution:

$$v = A_\alpha = (S_1 S_2) \begin{pmatrix} y \\ z \end{pmatrix} = S_{1_y} + S_{2_z} \tag{21}$$

which from Theorem 1 gives:

$$v = y + z \tag{22}$$

THEOREM 2: Let W be a topological vector space, and let W_1 and W_2 be complementary linear subspaces of W. Thus every vector v in W has a unique representation: $v = y + z$, with $y \in W_1$ and $z \in W_2$. The following statements are equivalent:

(a) W is the topological direct sum of W_1 and W_2;

(b) the mapping $y + z \mapsto y (y \in W_1,\ z \in W_2)$ is continuous;

(c) the mapping $y + z \mapsto z\ (y \in W_1,\ z \in W_2)$ is continuous.

PROOF: Define mappings u_1, u_2: $W \mapsto W$ as follows: if $v \in W$, write

$v = y + z$ with $y \in W_1$, $z \in W_2$ and define $u_1(v) = y$ and $u_2(v) = z$. By the uniqueness of such representations $v = y + z$, u_1 and v_2 are well defined and linear. Since $u_1(v) + u_2(v) = v$; for all v, clearly u_1 is continuous iff u_2 is continuous, thus (b), (c) are equivalent. If $T: W_1 \times W_2 \mapsto W$ is the mapping $T(y, z) = y + z$, then

$$T^{-1}(y+z) = (y, z) = (u_1(y+z), u_2(y+z)) \tag{23}$$

for all $y \in W_1$ and $z \in W_2$. That is, $T^{-1}(v) = (u_1(v), u_2(v))$ for all $v \in W$. Thus u_1, and u_2 are the compositions of T^{-1} with the coordinate projections of $W_1 \times W_2$. Therefore T^{-1} is continuous if and only if u_1 and u_2 are continuous; but the continuity of T^{-1} is, by definition, precisely the condition (a).

COROLLARY: If W is a topological vector space of dimension n, and W_1 is a linear subspace of W of dimension n, then for any $n \times n$ matrix, say Q, the transformation T which assigns to each vector v its product $Q\alpha$:

$$T : v \mapsto y \; Q\alpha \tag{24}$$

is a linear mapping; every component of W_1 is expressible as a homogeneous linear polynomial in the components of W.

Consider the topological vector space W in E^n of vectors $k = gx$. Let us assume, without any loss of generality, that the two topological complementary subspaces W_1 and W_2 of W have unitary bases. From the results in (23), the vector gx in W can be decomposed uniquely in the form

$$
\begin{aligned}
gx &= [\,(gx)^T + (1-gx)^T]^T \\
&= (xg)^T + (1-xg)^T \\
&= g^T x^T + (1-g^T x^T)
\end{aligned}
\tag{25}
$$

with $y \equiv g^T x^T \in W_1$ and $z \equiv 1 - g^T x^T \in W_2$.
The linear mapping T: $gx \mapsto g^T x^T$ which sends gx to its 'W_1 component' is called the *projection* of W onto W_1, along W_2. The corollary of this is that if $W = W_1 \oplus W_2$ is a decomposition of W into complementary subspaces, such that each k = gx in W can be expressed uniquely in the form v = y + z; with $y = g^T x^T \in W_1$ and $z = (1 - g^T x^T) \in W_2$. If we send each gx to its $g^T x^T$ component $g^T x^T$ in W_1 then this mapping is linear.

Let I represent the matrix of this projection. The above endomorphism is an idempotent element, that is $I^2 = I$.

DEFINITION: Let W be a topological vector space. A linear mapping I: W -> W is called a *projector* if $I^2 = I$.

THEOREM 3: Let W have the topological direct sum composition $W = W_1 \oplus W_2$. Let I be the matrix of the linear mapping T: $W \mapsto W$ such that I is a projector. Then

(i) $I\,(g^T x^T) = g^T x^T$ for all $g^T x^T \in W_1$;

(ii) $I\,(1 - g^T x^T) = 0$ for all $1 - g^T x^T \in W_2$

PROOF: Given that I denotes the transformation of gx to $g^T x^T$,

$$
I(gx) = g^T x^T \tag{26}
$$

Suppose we decompose $g^T x^T$ as we did gx. Then

$$g^T x^T = g^T x^T + 0 \tag{27}$$

with $g^T x^T \in W_1$ and $0 \in W_2$. Now

$$\begin{aligned} I(g^T x^T) &= I(I\,gx) = I^2(gx) \\ &= I(gx) \end{aligned} \tag{28}$$

since I is idempotent. It therefore follows that

$$I(g^T x^T) = g^T x^T \tag{29}$$

Consider the space W_2 with elements $1 - g^T x^T$. Since the spaces W is the topological direct sum of the spaces W_1 and W_2, it follows that $gx = g^T x^T + 1 - g^T x^T$. Hence

$$1 - g^T x^T = gx - g^T x^T \tag{30}$$

Therefore

$$I(gx - g^T x^T) = I(gx) - I(g^T x^T) \tag{31}$$

But $I(gx) = g^T x^T$ and $I(g^T x^T) = g^T x^T$. Therefore

$$I(1 - g^T x^T) = 0 \tag{32}$$

Hence if I is an idempotent endomorphism of W, then the subspace $W_1 = I(W)$ is mapped identically onto itself by I. Similarly, the set of all $gx - g^T x^T$, which forms a subspace W_2 of W is sent to zero by I. The decomposition

$$gx = I(gx) + (gx - I(gx)) = g^T x^T + 1 - g^T x^T \tag{33}$$

is unique. Since

$$gx = I(gx) + 1 - g^T x^T \tag{34}$$

it follows that

$$1 - g^T x^T = J(gx) \qquad where \; J = 1 - I \tag{35}$$

Hence J = 1-I is the projection of W onto W_2, with null space W_1, and

$$I(1 - I) = (1 - I)I = 0 \tag{36}$$

Thus we have established:

To every complementary decomposition $W = W_1 \oplus W_2$, there correspond two projections I and J, with $W_1 = I(W)$, $W_2 = J(W)$, $I\,J = J\,I = 0$, and $I + J = 1$.

We have seen earlier that the expansion derivative of g is given by $\dot{g} = - g$. Consider a small change in income of the amount τ. This gives a change in g, a differential of the form

$$dg = (d\tau)\,\dot{g} = -(d\tau)\,g \tag{37}$$

where $d\tau$ is any differential at a point g in the budget space. Since $x = f(g)$, and $gx = 1$, we get

$$d\tau = -(dg)x \tag{38}$$

For $d\tau = 0$, which is the absence of any income changes, $-(dg)\,x$ therefore defines the *substitution differential*.

Any budget differential dg at a point g is decomposible in a unique way into a sum of an expansion and a substitution:

$$dg = (d\tau)\,\dot{g} + dg^{*} \tag{39}$$

where dg^* is the substitution term; $(dg^*)x = 0$. Multiply the above equation by x, and we get

$$(dg)\, x \;=\; -(d\tau) \tag{40}$$

Therefore $d\tau = -(dg)\ x$, and hence (38) can be written as

$$(d\tau)\,\dot{g} \;=\; -((dg)\,x)\,(-g) \;=\; dg\,x\,g \tag{41}$$

Consider the term dg^*. From (39) and (41), we get

$$\begin{aligned} dg^* &= dg - (dg)\,xg \\ &= dg(1 - xg) \end{aligned} \tag{42}$$

Hence the decomposition in 39 can be restated as

$$dg \;=\; (dg)\,Xg + (dg)\,(1 - Xg) \tag{43}$$

by the transposition of which, we get

$$\begin{aligned} dg^T &= (g^T x^T)\,dg^T + (1 - g^T x^T)\,dg^T \\ &= I\,dg^T + J\,dg^T \end{aligned} \tag{44}$$

The matrices $I = g^T\ x^T$ and $J = 1 - g^T\ x^T$ have been shown earlier to be idempotent, and hence projectors.

Consider the Marshallian demand function $x = x(p,m)$. The matrix (S) of the Slutsky coefficients is given by

$$S \;=\; m(x_p + x_m x^T) \tag{45}$$

We have shown that $X_p\ m = X_g$ and $x_m\ m = -X_g\ g^T$. Hence the Slutsky matrix (S^*) corresponding to the normal demand function $x = f(g)$ is written as

$$S^* \;=\; x_g\,(1 - g^T x^T) \tag{46}$$

from which it follows that the Slutsky matrix is the product of a Jacobian and a projector.

The decomposition of the budget differential can be translated into a similar decomposition in the commodity space through the Slutsky matrix. Since $x = f(g)$, a budget differential dg leads to a corresponding commodity differential

$$dx = x_g (I\, dg^T + J\, dg^T) \tag{47}$$

where $I = g^T x^T$ and $J = 1 - g^T x^T$. Hence

$$\begin{aligned} dx &= x_g\, g^T (x^T dg^T) + x_g (1 - g^T x^T)\, dg^T \\ &= x_g\, g^T (dg\, x) + x_g (1 - g^T x^T)\, dg^T \end{aligned} \tag{48}$$

Now $x_g\, g^T = -\dot{x}$, $dg\, x = -d\tau$, and $X_g(1 - g^T x^T) = S^*$. Thus

$$dx = \dot{x}\, d\tau + S^*\, dg^T \tag{49}$$

Consider the term S^* dgT. It can be shown that $g\, S^*$ equals zero,

$$\begin{aligned} gS^* &= g\, x_g(1 - g^T x^T) \\ &= -x^T (1 - g^T x^T) \\ &= -x^T + (x^T g^T)\, x^T \\ &= -x^T + x^T \\ &= 0 \end{aligned} \tag{50}$$

The first term on the right hand side of (50) is the expansion term, and the second term is the substitution term, such that $g\, S^*$ dgT = 0.

We have the following theorem due to Afriat (1980):

THEOREM 4: For a differentiable demand function $x = f(g)$, and for any budget differential dg, the corresponding commodity differential dx has a unique resolution into a sum

$$dx = \dot{x}\, d\tau + S^*\, dg^T \tag{51}$$

where the first term is the income effect; $d\tau = -dg\, x$, and the second term the substitution effect, such that $g\, S^*\, dg^T = o$, and S^* is the Slutsky matrix.

1.2 The Classical Frobenius Theorem

The integrability question is concerned with the problem of providing conditions that assure that the demand functions can be integrated to yield the generating utility function. These conditions are obtained from a standard theorem of differential equations, as related to Frobenius. The implications of this theorem for the integrability of direct demand functions in the budget space can be expressed in terms of certain restrictions on the Slutsky matrix. This allows the direct demand functions to be integrated in the budget space to yield an integral surface, which is the income compensation function.

Although the integrability conditions in terms of the Slutsky matrix is well known, the link between the Frobenius theorem in its classical form and the condition of symmetry of the Slutsky matrix has never been shown. In this chapter, we explain this link, in terms of the conditions for the existence of an integral manifold of a system of differential equations, the latter being described as a *distribution* on a manifold. This condition, as given by the Frobenius theorem, requires the distribution to be *involutive*. We will show how this translates into the familiar restriction of symmetry of the Slutsky matrix.

The Problem Defined:

If m denotes the consumer's income; p the vector of prices, then the set of prices p and m together define the *budget space* Ω. The indirect utility function has been defined earlier as the adjoint of the direct utility function U:

$$V(p,m) = \sup \{ U(x);\ p^T x \le m,\ \ x \ge 0 \} \tag{1}$$

If $\Re$ denotes the set of reals then the function V is a mapping

$$V :\ \Omega \in E_{+}^{n+1} \mapsto \Re \tag{2}$$

that associates to each element (p,m) $\in \Omega$, an element V of $\Re$.

Consider the *income compensation function* developed by McKenzie (1957). It is defined on Ω as

$$\begin{aligned} \mu(p^r;\ p^*,m^*) &= \inf \{ m > 0;\ (p^r,m)\ R^*\ (p^*,m^*) \} \\ &= \inf \{ m > 0;\ V(p^r,m) \ge V^* \} \end{aligned} \tag{3}$$

where R^* denotes the indirect preference ordering, and p^r denotes some reference price (see Hurwicz and Uzawa, 1971). The function μ is a mapping

$$\mu :\ \Omega_n \times M \mapsto \Re \tag{4}$$

where M = { m $\in E^1$; such that for any p $\in \Omega_n$ (p,m) R^* (p^*,m^*) }. The function μ satisfies the following

$$\begin{aligned} \mu(p^*;\ p^*,m^*) &= m^* \\ \frac{\partial\mu(p;\ p^*,m^*)}{\partial p_i} &= x_i(p,\ \mu(p{:}\ p,m)) \qquad i = 1.....n \end{aligned} \tag{5}$$

The income compensation function is a modified expenditure function, and derives its properties from the latter. Consider the following lemma by Hurwicz and Uzawa (1971):

LEMMA: If R^* is continuous, then the function $V_p{}^*$ defined for any fixed p^* by:

$$V_{p^*}(p,m) = \mu(p^*;\ p,m) \tag{6}$$

is an indirect utility function representing R^*. Given the one-to-one correspondence between the direct and the indirect utility functions, the function μ can be used to construct, on the range Ω^n of the demand function, the utility function U(x):

$$U(x) \;=\; U^{p^*}(x) \;=\; \mu(p^*;\, p,m) \tag{7}$$

for any pair (p,m) in Ω, for which x = x(p,m).

The problem of integrability of the direct demand functions in the budget space is therefore one of finding the conditions under which the demand functions can be used to construct the income compensation function, such that the conditions in (6) are satisfied. Once the income compensation function has been shown to exist, the utility function can be derived from it.

Given that the functions V and μ are inverses of each other, we can define a function χ as follows:

$$\begin{aligned} \chi(\mu)\,[p,m] &= Max\ \{\ V;\ m \geq \mu(p;\ p,m)\ \} \\ \chi^{-1}(V)\,[p,V^*] &= Min\ \{\ m;\ V(p,m) \geq V^*\ \} \end{aligned} \tag{8}$$

where the function χ is a bijection with an inverse χ^{-1} between the class of income compensation functions 1 and the indirect utility functions. Hence the function χ is homeomorphic, or the sets Ω and $\Re$ are homeomorphic to each other.

Consider the domain (Ω_n x M) of the income compensation function μ. For notational convenience, let us denote this topological space by N. We can define a local chart of dimension n+1 on N as a pair (G, γ), where G is an open set in N and γ: G $\mapsto E_+^{n+1}$ is a homeomorphism from G to an open set γ(G) in E_+^{n+1}.

An atlas $\mathbb{X}$ of dimension n+1 is a collection of local charts whose domains cover N and such that if (G,γ), $(G',\gamma') \in \mathbb{X}$ and $G \cap G' \neq \emptyset$, then the map $\gamma' \circ \gamma^{-1}$: $\gamma(G \cap G') \mapsto \gamma'(G \cap G')$ is a diffeomorphism between open sets of E_+^{n+1}.

For every integer i, let

$$\beta_i : E^{n+1} \mapsto \Re \tag{9}$$

denote the natural projection. β^i denotes the canonical coordinates on E^{n+1} which maps a point $e = (e^1, \ldots, e^{n+1}) \in E^{n+1}$ into $\Re$ by

$$\beta_i(e_1 \ldots\ldots e_{n+1}) = e_i \tag{10}$$

Let $\beta_i \circ \gamma$ be abbreviated by α^i (where α: $I = [a,b] \mapsto E^{n+1}$). The composite mappings $\beta^i \cdot \gamma$: $G \mapsto \Re \equiv \gamma^i$ is the ith coordinate function in G with respect to γ, and maps N into $\Re$ by $\alpha \mapsto \alpha^i$. The coordinates $(\alpha^1, \ldots, \alpha^{n+1}) \equiv (p_1, \ldots, p_n, m)$ of the image $\gamma(G) \in E^{n+1}$ are the coordinates with respect to the chart, and hence form a local coordinate system.

Consider the domain N of the income compensation function μ. We can say that N is an n+1 dimensional *manifold* since for every point in N there is an open neighborhood which is homeomorphic to an open set in E^{n+1}. This is motivated by the fact that for every point in N, there is an n+1 dimensional chart, with coordinate functions defined by $\alpha = (\alpha^1 \ldots \alpha^{n+1}) = (p_1 \ldots p_n, m)$.

Throughout this chapter, we are interested in differential equations which are defined in an open subset of the space N. The underlying geometric features of this differential equation are independent of any particular coordinate system on the subset which may be used to define it, although the functions V and μ are defined with respect to a coordinate system.

Let q denote any given point of the n+1 dimensional manifold N. Let $\zeta(N,q)$ denote the set of all local smooth functions at the point q of N. If the set $\zeta(N,q)$ permits the operations of addition and multiplication, as well as scalar multiplication for any arbitrary elements in $\zeta(N,q)$, then it defines an algebra over the field $\Re$ of real numbers which is associative and commutative.

It should be pointed out that the operations of scalar multiplication, addition and multiplication, are not adequate by themselves to make any set of smooth functions, say $\mathfrak{v}(N,q)$ an algebra over $\Re$. In fact $\mathfrak{v}(N,q)$ is not even a linear space over $\Re$, since for any two smooth functions $f_1: B_1 \mapsto \Re$ and $f_2: B_2 \mapsto \Re$, we have

$$f_1 + (-f_1) \neq f_2 + (-f_2)$$

unless $B_1 = B_2$. To remove this undesirable behavior of $\mathfrak{v}(N,q)$, let us define a relation ~ in the set $\mathfrak{v}(N,q)$ as follows. For any two local smooth functions $f_1 : B_1 \mapsto \Re$ and $f_2 : B_2 \mapsto \Re$ in $\mathfrak{v}(N,q)$, we have $f_1 \sim f_2$ iff there exists an open neighborhood $W \subset B_1 \cap B_2$ of q in N such that $f_1(w) = f_2(w)$ for every w in W. Since this relation in $\mathfrak{v}(N,q)$ is reflexive, symmetric, and transitive, it is an equivalence relation. Hence ~ divides the members of $\mathfrak{v}(N,q)$ into disjoint equivalence classes called the *germs of local* *smooth functions* at the point q of N. Let

$$\zeta(N,q) \quad = \quad \vartheta(N,q) \;/\; \sim$$

denote the set of all smooth germs at the point q of N, and let

$$\beta \;:\; \vartheta(N,q) \mapsto \zeta(N,q)$$

be the natural projection of the set $\vartheta(N,q)$ into its quotient set $\zeta(N,q)$. From the definition of the relation ~ in $\zeta(N,q)$, if $f_1 \sim f_2$ and $h_1 \sim h_2$ in $\zeta(N,q)$, then

$$af_1 + bh_1 \;\sim\; af_2 + bh_2, \qquad f_1\, h_1 \;\sim\; f_2\, h_2$$

for every a, b in $\Re$. The operations of scalar multiplication, addition and multiplication in $\zeta(N,q)$ now define an algebra over $\Re$, which is obviously associative and commutative.

Given the class $\zeta(N,q)$ of all smooth real valued functions, we can define a *tangent vector* as a derivation on the algebra of germs of differentaible functions at $q \in N$. A tangent vector φ is a mapping of $\zeta(N,q)$ into the field of real numbers:

$$\varphi \;:\; \zeta(N,q) \;\mapsto\; \Re \tag{11}$$

which satisfies the following two conditions:

(i) φ is a linear functional on the space $\zeta(N,q)$;

that is, for any f_1, $f_2 \in \zeta(N,q)$, and $a,b \in \Re$ we have:

$$\varphi(af_1 + bf_2) \;=\; a\,\varphi(f_1) \;+\; b\,\varphi(f_2)$$

(ii) φ is a *derivation* that is for any f_1, $f_2 \in \zeta(N,q)$

$$\varphi(f_1 f_2) \;=\; \varphi(f_1)\, f_2(q) \;+\; f_1(q)\, \varphi(f_2)$$

Let us denote the set of all tangent vectors of N at the point qby T(N,q). For any v, w in T(N,q) and a, b in $\Re$, the function

$$av + bw : \zeta(N,q) \mapsto \Re$$

defined by

$$(av + bw)(\omega) = av(\omega) + bw(\omega)$$

for every $\omega \in \zeta(N,q)$ is in $T(N,q)$. This makes $T(N,q)$ a linear space over the field of $\Re$, called the *tangent space* of N at the point q.

If $\gamma : G \mapsto E^{n+1}$ is a local chart, and $\{e^1,\dots e^{n+1}\}$ is a canonical basis for E^{n+1}, then the set $\{[\gamma^{-1}(te^1)],\dots,[\gamma^{-1}(te^{n+1})]\} = \partial/\alpha^1,\dots,\partial/\partial\alpha^{n+1}$ is a basis for $T(N,q)$, te^i being the curve $t \mapsto te^i$, $t \in \Re$. The space $T(n,q)$ has a natural structure of a real vector space, and its dimension is the dimension of N, that is, n+1. Hence the chart (G,γ) induces an isomorphism of $T(N,q)$ onto E^{n+1}.

Let $q \in N$, (G,γ) be a coordinate system with $q \in G$, and γ^i α^i, $i = 1\dots n+1$. For a function f in $\zeta(N,q)$, we can define, for each i, a tangent vector φ^i in $T(N,q)$

$$\varphi^i = \left(\frac{\partial}{\partial\alpha^i}\right)_q f = \frac{\partial(f\circ\gamma^{-1})}{\partial\alpha^i}(\gamma(q)), \quad f \in \zeta(N,q) \tag{12}$$

It can be easily checked that the tangent vectors φ^i at the point $q \in N$ is a linear derivation of the algebra $\zeta(N,q)$.

Given that $\alpha = \alpha^1\dots\alpha^{n+1}$ is the coordinate system at q, then for every integer $i = 1\dots n+1$, the operators

$$\varphi^i = \left(\frac{\partial}{\partial\alpha^i}\right)_q (f) = \left(\frac{\partial f}{\partial\alpha^i}\right)_q \tag{13}$$

define the components of the tangent vector φ with respect to the local coordinate system $\alpha^1 \ldots \alpha^{n+1}$. We therefore have

$$\varphi = \sum \frac{\partial f}{\partial \alpha^i} \qquad (14)$$

We interpret (14) as the directional derivatives of f at q, in the α^i coordinate direction.

Vector Analysis on Manifolds

Consider the differentiable manifold N, and T(N,q) its tangent space. A map Y on a subset Q of the manifold N which assigns to each q in N, a tangent vector $\varphi(q) \in T(N,q)$ at q by the mapping

$$Y: N \mapsto \varphi_q \qquad (15)$$

is a *vector field* on N. The vector field Y is also referred to as an infinitesmal transformation, a mapping which assigns to every point q in N, a tangent vector Y(q) to N at this point.

For any manifold N and its tangent space T(N,q) there is a natural projection π: T(N,q) $\mapsto$ N, which maps T(N,q) on to the point q. A vector field Y is a map Y: N $\mapsto$ T(N,q) such that $\pi \cdot Y$ = identity on N. Geometrically a vector field describes a cross section of the projection of T(N,q) onto the manifold.

Derivations and Lie Algebra

Let us denote the set of all vector fields on N by $\hbar(N)$. Let $C^2(N)$ denote the set of all continous functions of class C^2. If f is a function in $C^2(N)$,

$$f \in C^2, \quad q \mapsto Y(q)(f) \tag{16}$$

it defines another function on N which also belongs to $C^2(N)$. Y is therefore a linear mapping

$$C^2(N) \mapsto C^2(N) \tag{17}$$

called the *Lie derivative* of f by Y, and Y is a derivation of C^2. We will see below that in terms of local coordinates, a derivation is nothing but a first order differential operator.

For any Y', $Y^2 \in \hbar(N)$, and f_1, f_2 in C^2, the correspondences

$$q \mapsto f(q)\ Y_q; \quad and \quad q \mapsto Y_q^1 + Y_q^2 \tag{18}$$

again define vector fields on N. These vector fields denoted by fY and $Y' + Y^2$ respectively, and are the product of f and Y, and the sum of Y' and Y^2. The rules for products and sums are:

$$\begin{aligned} (a) \qquad f_1(f_2Y) &= (f_1f_2)Y \\ (b) \qquad f(Y^1 + Y^2) &= fY^1 + fY^2 \\ (f_1 + f_2)Y &= f_1Y + f_2Y \end{aligned} \tag{19}$$

Consider the set Q defined as an open subset of the manifold N, with $\alpha^1 \ldots \alpha^{n+1} \equiv p_1 \ldots p_{n+1}$, defined as a local coordinate system for the manifold N. Let $Y \in \hbar(N)$ assign a tangent vector to each q in N. Setting

$$Y^1(q)\mu = \frac{\partial\mu}{\partial\alpha^1} \equiv \frac{\partial\mu}{\partial p_i} \tag{20}$$

we obtain a tangent vector at q, and the mapping $q \mapsto Y_1(q)$ is a vector field defined on the open set Q. If we set

$$Y_i(q) = \frac{\partial l}{\partial \alpha^i} \equiv \frac{\partial l}{\partial p_i} \qquad (21)$$

we obtain n vector fields which are linearly independent.

If μ is the income compensation function, and Y_i and Y_j any two vector fields defined on the manifold N, the operation $Y_i \cdot Y_j$ is not a vector field, since as an operator, it is a second order partial differential operator. Since $Y_i(\mu) = \partial\mu/\partial p_i$ and $Y_j(\mu) = \partial\mu/\partial p_j$, we have $Y_i Y_j(\mu) = \partial^2\mu/\partial p_i \, \partial p_j$, which is not a tangent vector. However the operation $Y_i Y_j - Y_j Y_i$ is a derivation, since

$$[Y_i Y_j - Y_j Y_i](\mu) = Y_i(Y_j(\mu)) - Y_j(Y_i(\mu)) \qquad (22)$$

cancels out the middle terms: $Y_i Y_j(\mu)$ and $Y_j Y_i(\mu)$. Hence $Y_i Y_j - Y_j Y_i$ defines a vector field, called the *Jacobi* or *Lie Bracket* of Y_i Y_j, and is denoted by $[Y_i \; Y_j]$.

This bracket operation, which assigns to every pair of vector fields $(Y_i \; Y_j)$ the infinitesmal transformation $[Y_i \; Y_j]$, is a law of composition for the vector fields. If Y_i, Y_j, Y_k are three vector fields on N, the following formal laws follow from direct computation:

(1) The Lie bracket is *additive* in each variable:

$$[Y_i + Y_j, Y_k] = [Y_i, Y_k] + [Y_j, Y_k]$$
$$[Y_i, Y_j + Y_k] = [Y_i, Y_j] + [Y_i, Y_k]$$

(2) The Lie bracket is *skew symmetric*:

$$[Y_i, Y_j] = -[Y_j, Y_i]$$

hence $\qquad [Y_i \; Y_i] = 0$

(3) The Lie bracket is *not associative:*

$$[[Y_i,Y_j],Y_k] \neq [Y_i,[Y_j,Y_k]]$$

but satisfies the following:

$$[[Y_i,Y_j],Y_k] + [[Y_j,Y_k],Y_i] + [[Y_k,Y_i],Y_j] = 0$$

which is referred to as the *jacobi identity*.

Submanifolds and Distributions:

We now address the system of partial differential equations

$$\frac{\partial m}{\partial p_i} = x_i(p,m(.)) \qquad i = 1....n \tag{23}$$

where $m = \mu(p)$ is the unknown function of p, with values in E^1, and the given function $x(p,m)$ is a mapping $x: \Omega_n \times E^1 \mapsto \Omega^n$ of class C^1.

In the function $\mu(p;p,m)$ fix p at p^*. From (6), let

$$V_p^* = \mu(p^*;p,m) \tag{24}$$

be the solution equal to V^* for $p = p^*$, and we have

$$V^* = V(p,m) \tag{25}$$

where V is of class C^1. Thus the function $V(p,m)$ is defined on E^1, and is constant on every solution $\mu = \mu(p)$, supposed completely integrable.

The problem of finding the integrability conditions on x is therefore equivalent to the solution of the system of partial differential equations in (23).

Given an n+1 dimensional manifold N in which the function μ is defined, it has $(p_1...p_n, m)$ as a coordinate system. Consider the subspaces spanned by the function Y_i. Let us assign to every q in the manifold N, an n dimensional subspace D(q) of N at q. A function D which assigns to each q in N an n dimensional subspace $D(q) \subset T(N,q)$ is called an n dimensional *distribution*. A distribution is smooth (of class C^1), if for every $q \in N$, there is a neighborhood of G of q, in which there are C^1 vector fields $Y_1 \ldots Y_n$, such that for s in N, $Y_1(s) \ldots Y_n(s)$ is a local basis of D(s). The Y_j, $j = 1 \ldots n$ are said to generate D on G.

We now wish to relate the n dimensional distribution D(q), with a coordinate system $\alpha^1 \ldots \alpha^n$, to the n+1 dimensional manifold N with $\alpha^1 \ldots \alpha^{n+1}$ as the coordinate system.

Let Z be a subset of N, such that each $z \in Z$ is in the domain of a chart (G,γ) of N such that

$$\gamma : G \cap Z \mapsto E^n \times E \text{ by } \gamma(p) = (p_1, \ldots, p_n, m) \tag{26}$$

Then Z is a submanifold of N. This submanifold Z has to be placed in the containing manifold N in quite a special way. To describe the special relationship between the submanifold Z and N, we introduce the notion of a *coordinate slice*. We define an n dimensional slice in a manifold N as a set of points in a coordinate neighborhood with n+1 coordinates $(p_1 \ldots p_n, m)$ of the form:

$$\{ q; \;\; q \in G, \;\; \alpha^{n+1} = m = c \} \tag{27}$$

where the c is a constant determining that slice. A coordinate slice is the image under the inverse of a coordinate map of that part of the n dimensional plane which lies in the coordinate range.

PROPOSITION 1: If Z is a submanifold of N, then for every $q \in N$, there are coordinates $\alpha^1 \ldots \alpha^{n+1}$ for N in a neighborhood G of q such that the coordinate slice corresponding to constants $\alpha^{n+1} = c$ is a neighborhood of $q \in Z$. Further, the restriction of $\alpha^1 \ldots \alpha^n$ to that slice are coordinates of Z.

The submanifold Z is an *integral manifold* of the n dimensional distribution D(q), if for every $z \in Z$, the tangent space of z, T(Z,z) is contained in the distribution D(z).

DEFINITION: The distribution D is completely integrable, if for any $q \in G$ in N, there is a coordinate system $\alpha = \alpha^1 \ldots \alpha^{n+1}$, and

$$Z = \{ \alpha \in G; \ \alpha^{n+1} = c \} \tag{28}$$

such that the submanifolds given by Z are integrals of D.

Consider an n dimensional subspace of the manifold N. The vector field Y, defined as $Y_i = \partial / \partial p_i$ for $i = 1 \ldots n$, belongs to the distribution D, since for every q in N, $Y(q) \in D(q)$. We can say that the distribution D is *involutive,* if for any $q \in N$ there is a neighborhood G and vector fields $Y^1 \ldots Y^n$ generating D, such that we have

$$[Y_i \ Y_j] \in D \quad for \ i,j = 1 \ldots . n \tag{29}$$

Note that $Y_i(\mu) = \partial / \partial p_i \ (\mu) = x_i > 0$. The interpretation of (29) is that since Y is a vector field which is nonvanishing the hypersurface D through q, orthogonal to $\mu(p;p,m)$, is tangent at q to the indifference hypersurface of q, and μ indicates a direction of preference.

A one dimensional distribution in a plane which corresponds to some nonvanishing vector field is always integrable, since the local basis field will always have integral curves. In fact, a one dimensional distribution possesses a family of integral curves, which reduces to a field of parallel lines by a diffeomorphism (see Katzner 1970 for an intuitive proof). However, in a three or more dimensional space, a distribution whose dimension is greater that one, need not have an integral surfaces at all. For example, the vector fields ∂_x, and $\partial_y + \partial_z$ in a 3-dimensional space $\Re^3$, span a 2-dimensional distribution but $[\partial_x, \partial_y + x\partial_z] = \partial_z$ does not belong to the distribution. The problem of characterizing those distributions that do posess integral surfaces, is the integrability problem.

Frobenius Theorem

The theorem of Frobenius is concerned with the problem of determining integral submanifolds of vector fields, with the property that each vector field is tangent to the submanifold at that point.

THEOREM: Let D be an n dimensional smooth distribution defined in the manifold N. If D is involutive, that is

$$Y_i,\ Y_j \in D \rightarrow [Y_i,\ Y_j] \in D \tag{30}$$

then D is completely integrable. Locally there are coordinates α^i for i = 1...n, such that $\partial/\partial\alpha^i$ are a local basis for D, and the coordinate slice α^{n+1} is an interal submanifold, and $\alpha^{n+1} = m$ is the solution function.

We now wish to apply the Frobenius theorem to the system of partial differential equations:

$$\frac{\partial m}{\partial p_i} = x_i(p,m) \qquad i = 1....n \tag{31}$$

where $m = \mu(p)$ is a function in N, with values in E^1, such that

(i) $(p,\mu(p)) \in N$, for every $p \in \Omega_n$,

(ii) $\mu'(p) = x(p,\mu(p))$, for every $p \in \Omega_n$,

is called a solution of equation (31).

The n dimensional subspace D of x is a distribution. The operators $Y_i = \partial/\partial p_i$ form a vector field on Z. Since $\partial\mu/\partial p_i = x_i$ hence $Y_i..Y_n$ are independent vector fields spanning D. Therefore the vector fields $Y_i(\mu)$ belong to D, and D is a smooth distribution.

Let us now consider the Lie bracket $[Y_i, Y_j]$. Given the definition of the bracket $[Y_i, Y_j](\mu) = Y_i, Y_j\ (\mu) - Y_j, Y_i\ (\mu)$ we have:

$$[Y_i,\ Y_j]\,(\mu) = \frac{\partial}{\partial p_i}\left(\frac{\partial\mu}{\partial p_j}\right) - \frac{\partial}{\partial p_j}\left(\frac{\partial\mu}{\partial p_i}\right) \tag{32}$$

Since $\partial\mu/\partial p_i = x_i(P,\mu(\cdot))$, we have:

$$[Y_i,\ Y_j]\,(\mu) = \frac{\partial x_j(p,\ \mu(\cdot))}{\partial p_i} - \frac{\partial x_i(p,\ \mu(\cdot))}{\partial p_j} \tag{33}$$

By the theorem of Frobenius, the distribution D is completely integrable, if the Lie bracket in (34) also belongs to D, which would imply that D is involutive.

PROPOSITION: Let $Y_1 \ldots Y_n$ be vector fields on Z which are linearly independent at every point z in Z. If $[Y_i, Y_j] = 0$, then for any z in Z, we have a coordinate system such that $Y_1 \ldots Y_n = \partial/\partial p_1 \ldots \partial/\partial p_n$ is the associated vector fields. This implies that Y(z) span the tangent space T(Z,z) and therefore generate D(q) for $q \in G$. Hence we can say that if

$$if \;\; [Y_i,\, Y_j] = 0 \;\; \rightarrow \;\; [Y_i,\, Y_j] \in D(q) \tag{34}$$

and the distribution D(q) is involutive.

If the bracket $[Y_i, Y_j]$ equals zero, then

$$\frac{\partial h_i(P,U)}{\partial p_j} = \frac{\partial h_j(P,U)}{\partial p_i} \qquad i,j, = 1 \ldots n \tag{35}$$

which implies the symmetry of the Slutsky matrix S(p,m).

Equation (35) defines the sufficient conditions for the existence of integral surfaces, which is stated in terms of the existence of a solution to the system of differential equations in (31). This solution is unique, that is there is only one indifference surface through some initial conditions $m^\circ = p^\circ x^\circ$, provided the demand functions x(p,m) are of class C^1; that is, differentiable everywhere. However, this differentiability condition can be weakened to require that the demand functions merely be *Lipschitzian* in m.

DEFINITION: The function x(p,m) satisfies a Lipschitz condition in m, if there exists a constant $K > 0$ such that:

$$| x(p, m^1) - x(p, m^2) | \leq K | m^1 - m^2 |$$

for each m^1 m^2 in E.

The above symmetry conditions are sufficient for the existence of integral surfaces. For these to be interpreted as the indifference surfaces corresponding to the generating utility function, these integral surfaces have to be convex, such that quasiconcavity of the utility function is guaranteed. This imposes an additional restriction on the Slutsky matrix: the condition of negative semidefinite. (Hurwicz, 1971; and Hurwicz and Uzawa, 1971; Afriat, 1980)

2.3 Integrability Revisited: The Dual Form Of The Frobenius' Theorem

In section 2.2, the integrability question was considered in terms of the conditions on the demand functions that would yield hypersurfaces in the price space which corresponds to the income compensation function. This problem was formulated in terms of the solution to a system of partial differential equations. The Frobenius' theorem in its classical version characterizes the integrability of a distribution by a condition on the Lie bracket of the generators of the distribution. An alternative approach to the integrability problem is considered here: the integration of a system of first order partial differential equations reduces to the integration of a system of ordinary differential equations. The basis of this reduction is a geometric analysis of the formation of surfaces from families of curves. We begin with these geometric arguments and then apply them to the partial differential equations.

To do this we use the notion of *differential forms* which play a key role in the topological aspects of differential geometry. This approach uses Grassmann or exterior algebra, composed of functions called alternate multilinear functions with the Grassmann multiplication as the law of composition. We use this algebra to define the exterior differential forms of Cartan on a manifold. The special nature of alternate mapping results in certain unique properties for the multiplication as well as the differentiation of these forms. We use these results to characterize the integrability of the distribution, and this constitutes the dual form of the Frobenius' theorem.

Consider the income compensation function

$$\mu(p^r;\ p^*,m^*) \ = \ inf\ [\ m > 0;\ (p^r,m)\ R^*\ (p^*,m^*)\ ] \qquad (1)$$

The domain of μ, denoted by N, is an n+1 dimensional manifold. In the last section, the integrability of a distribution D was defined in terms of the solution to a system of partial differential equations:

$$\frac{\partial m}{\partial p_i} \ = \ x_i(p,m) \qquad i = 1\ldots n \qquad (2)$$

where $x_i = Y_i$ was shown to be a vector field on N, defined as a mapping which assigns to every point q ϵ N, a tangent vector to N at q.

The Differential of the function $\mu(P;\ P,M)$

Let q denote a point in the n+1 dimensional manifold N. If $\mathfrak{F}(N,q)$ denotes the set of all local smooth functions at the point q of N, then $\zeta(N,q) = \mathfrak{F}(N,q)\ /\ \sim$ denotes the set of all smooth germs at the point q of N, consisting of the equivalence class of local smooth functions at the point q ϵ N (see chapter 2.2). The algebra of the linear space $\zeta(N,q)$ is associative and commutative.

Consider a subset K(N,q) of $\zeta(N,q)$, which consists of all smooth germs $\omega \in \zeta(N,q)$, such that $\varphi(\omega) = 0$ holds for every tangent vector $\varphi \in T(N,q)$, where T(N,q) denotes the union of all tangent spaces to the manifold N. If $\zeta(N,q)$ is an algebra, then $K(N,q)$ is a subalgebra of the algebra $\zeta(N,q)$. In particular, $K(N,q)$ is a linear subspace of the space $\zeta(N,q)$. The quotient linear space

$$T^*(N,q) \quad = \quad \zeta(N,q)\,/\,K(N,q) \tag{3}$$

is the *cotangent* space of N at the point q.

Consider the composition

$$\Phi \quad = \quad \pi^1 \circ \pi^2 \;:\; \vartheta(N,q) \;\mapsto\; T^*(N,q) \tag{4}$$

of the natural projections π^1: $\vartheta(N,q) \mapsto \zeta(N,q)$ and π^2: $\zeta(N,q) \mapsto T^*(N,q)$. The n+1 dimensional space $T^*(N,q)$ is the dual of the tangent space $T(N,q)$ and is the collection of all linear real valued functions on the vector space. $T^*(N,q)$ is called the cotangent space at q, or the space of covariant vectors.

In chapter 2.2, we considered the system of partial differential equation

$$\frac{\partial m}{\partial p_i} \quad = \quad x_i(p,m) \qquad i = 1,\ldots,n \tag{5}$$

where the n functions x_i are given functions of class C^1 defined in Ω. We required that, for (p,m) in Ω, this equation possess a solution $m = \mu(p)$, with the initial value $\mu(p^0) = m^0$.

Consider the differential of the function μ at the point q

$$(d\mu)_q \quad = \quad \varphi\mu|_q \tag{6}$$

which in term of the local coordinates is

$$(d\mu)_q = \sum_{i=1}^{n} \left\{ \left(\frac{\partial}{\partial p_j} \right)_q \mu \right\} (dp_j)_q \tag{7}$$

We define a cotangent vector Φ in $T^*(N,q)$ as

$$\Phi = dm - \sum_{i=1}^{n} x_i(p,m)\ dp_i \tag{8}$$

The cotangent space $T^*(N,q)$ of the n+1 dimensional manifold N at an arbitrary point q of N is an n+1 dimensional linear space over $\Re$. For the smooth chart α: $\alpha^1 \ldots \alpha^{n+1}$, $d\alpha^1 \ldots d\alpha^{n+1}$ = $(dp_1 \ldots dp_n, dm)$ form the basis of $T^*(N,q)$, dual to the basis of the tangent space $T(N,q)$. Since $\alpha^1 \ldots \alpha^{n+1}$ form a coordinate system at q, their differentials are linearly independent.

Multilinear Alternating Mappings

Let ϖ denote an arbitrary n dimensional linear space over the field $\Re$ of real numbers. For any positive integer r, let

$$\varpi^r = \varpi \oplus \ldots . \oplus \varpi \tag{9}$$

of r linear spaces, each of which is identical to ϖ.

An *r linear* function on ϖ is a mapping f:

$$f : \varpi^r \mapsto \Re \tag{10}$$

such that for every i = 1...r, f is linear in the ith variable, when the remaining r-1 arguments are kept fixed. Let us denote the space of the mapping of the r-linear function by $T^*{}_r$, where the subscript r denotes the degree of linearity.

The mapping Φ defined as the differential of the function μ, is therefore a 1-linear function; that is, Φ is a *linear mapping* of T into $\Re$. We denote the subspace of this mapping as T^*_1, or simply as T^*.

The mapping Φ is *alternating,* in the sense that it is skew symmetric. That is $\Phi(\alpha^1..\alpha^{n+1})$ changes signs whenever neighboring arguments are permuted. To explain this, we require a few preliminaries concerning permutation groups.

PERMUTATION GROUPS

Let Σm denote the group of all permutations of the set $\{1,..,m\}$ of the first m integers. We know that it contains m! elements. A *transposition* is a permutation $\sigma \in \Sigma m$, such that there exist integers i,j, for $i \neq j$, and $1 \leq i \leq m$, and $1 \leq j \leq m$, for which

$$\sigma(i) = j \quad and \quad \sigma(j) = i \qquad \textbf{(11)}$$

Note that σ interchanges i and j. It is evident that σ^2 is the identity element. It can be shown that any permutation $\sigma \in \Sigma m$, can be written as the product of transpositions, each of which interchanges two consecutive integers. The *signature* of a permutation σ, is denoted by $\epsilon(\sigma)$, and is equal to +1 if σ is the product of an even number of transpositions, and -1 if σ is the product of an odd number of transpositions.

We can now define an alternating mapping as follows:
Let $\Phi \in T^*$ be a multilinear alternating mapping. Then
(i) each time $\alpha^i = \alpha^j$ for a pair (i,j) of distinct suffixes:

$$\Phi(\alpha^1 \ldots . \alpha^{n+1}) = 0 \tag{12}$$

(ii) for each permutation σ of {1...n} we have

$$\Phi(\alpha_{\sigma(1)} \ldots . \alpha_{\sigma(n+1)}) = \epsilon(\sigma)\ \Phi(\alpha^1 \ldots .. \alpha^{n+1}) \tag{13}$$

where $\epsilon(\sigma) = \pm 1$, which is the signature of the permutation.

Given that Φ is a multilinear alternating mapping, it follows that it is skew symmetric or antisymmetric. Simply stated, it implies that Φ changes signs when neighboring arguments are permuted:

$$\Phi(\alpha^1, \alpha^2, \alpha^3, \ldots \alpha^{n+1}) = -\Phi(\alpha^2, \alpha^1, \alpha^3, \ldots \alpha^{n+1}) \tag{14}$$

Exterior Product

Let $\Phi_1(g)$ and $\Phi_2(h)$ be two alternating mappings defined in T^*, with $\Phi_1(g) \in T_r^*$ and $\Phi_2(h) \in T_s^*$. The product of Φ_1 and Φ_2 defines a function $\Phi_1 \wedge \Phi_2 : T_{r+s}^* \mapsto \Re$ as follows: If r = 0 or s = 0, $\Phi_1 \wedge \Phi_2$ is defined by scalar multiplication; otherwise $\Phi_1 \wedge \Phi_2$ is defines a new function $\Theta(g,h) \in T_{r+s}^*$

$$\Phi_1 \wedge \Phi_2 = \Sigma\ \epsilon(\sigma)\ \theta(\Phi_1(g), \Phi_2(h)) \tag{15}$$

where the summation is over all permutations σ of {1....g+h}, and it can be easily checked that there are (r + s)! / r! s! such permutations.

The assignment $\Theta : (\Phi_1, \Phi_2) \mapsto \Phi_1 \wedge \Phi_2$ is multilinear and continuous, which is referred to as *exterior product* and has the following properties:

(a) $\wedge$ is associative.

(b) $\wedge$ has $1 \in \Re\, T_0^*$ as a two sided unit; $1 \wedge \Phi = \Phi = \Phi \wedge 1$ holds for every element Φ of T^*.

(c) $\wedge$ is not commutative, but we have $\Phi_1 \wedge \Phi_2 = (-1)^{rs}\, \Phi_2 \wedge \Phi_1$ for every $\Phi_1 \in T_r^*$ and $\Phi_2 \in T_s^*$.

Hence the exterior multiplication $\wedge$ makes the linear space T^* an algebra over $\Re$. This algebra consisting of the alternate multilinear functions, with the exterior multiplication as the law of composition, is called *Grassmann algebra* of the linear space N. Because of the direct sum decomposition

$$T^*(N) = \sum_{k=0}^{n+1} T_k^*(N) \tag{16}$$

the Grassmann algebra is a graded algebra over $\Re$.

The Differential Forms of Cartan

Let ϖ be an arbitrary n+1 dimensional linear space over the field $\Re$ of real numbers. Then ϖ^r denote the direct sum of r linear spaces. A mapping f as an r linear function which assigns to every point in ϖ^r an element of order r in the subset T_r^*:

$$f : \varpi^r \mapsto T_r^* \tag{17}$$

is called a *differential form* of order r. In its canonical form, a differential form of degree r is

$$f = \sum f_{i_1 \ldots i_r}\, d\omega_1 \ldots\ldots d\omega_r \tag{18}$$

Consider the income compensation function $\mu(p;\, p,m)$, where the domain N of the function is an n+1 dimensional manifold. The Grassmann algebra associated with the space $T^*(N,q)$ is called the Cartan differential algebra at the point q in N.

We are interested in differential forms on N as a cross section of the vector bundle T^*; such that it assigns to each point q, a multilinear, skew symmetric form on the tangent vectors to N.

In addressing the integrability problem, differential forms of orders 0, 1, and 2 will be of interest. A 0-form is none other than the function μ itself. This is evident from the fact that μ is a mapping $\mu\colon N \mapsto T_0^* = \Re$. A differential 1-form is the alternating mapping $\Phi\colon N \mapsto T_1^* = T^*$. In its canonical form, the 1-form can be written as

$$\Phi = dm - \Sigma\, x_i(p,m)\; dp_i \tag{19}$$

A differential 2-form is a bilinear alternating mapping which assigns to each point q, an element of degree two in T_2^*:

$$d\Phi = \sum \Phi_{ij}\; d\alpha^i\, d\alpha^j \tag{20}$$

The multiplication of two forms is given by the exterior product of the forms. Let Φ_1 be an r-form and Φ_2 an s-form. The exterior product of Φ_1 and Φ_2 therefore defines the graded Grassmann algebra, and is associative and anticommutative. Given that Φ is a differential 1-form, we have

$$\Phi_1 \wedge \Phi_2 = - \Phi_2 \wedge \Phi_1 \tag{21}$$

Moreover, for every $\Phi_i \in T^*$,

$$\Phi_i \wedge \Phi_i = 0 \tag{22}$$

This follows from the fact that $\Phi_i \wedge \Phi_i = (-1) \Phi_i \wedge \Phi_i$, which implies $2(\Phi_i \wedge \Phi_i) = 0$, and hence $\Phi_i \wedge \Phi_i = 0$.

Consider two 1 forms: $\Phi_1 = \Sigma f_i d\alpha^i$, and $\Phi_2 = \Sigma g_j d\alpha^j$. The product of Φ_1 and Φ_2 can be written as:

$$\Phi_1 \wedge \Phi_2 = \Sigma f_i g_j \; d\alpha^i d\alpha^j$$

Now the term $d\alpha^i d\alpha^j$ can be viewed as the product of the 1-form $d\alpha^i$ with the 1-form $d\alpha^j$. Hence $d\alpha^i d\alpha^j$ is an exterior product and has the property:

$$d\alpha^i d\alpha^j = - d\alpha^j d\alpha^i \quad \text{and} \quad d\alpha^i d\alpha^i = 0$$

Hence

$$\begin{aligned} \Phi_i \wedge \Phi_j &= \sum_{1 \le i \le j \le n} (\Phi_i \Phi_j - \Phi_j \Phi_i) \; d\alpha_i \, d\alpha_j \\ &= \Phi_i \, d\alpha_i - \Phi_j \, d\alpha_j \end{aligned} \tag{23}$$

Exterior Derivatives

We now consider an important operation which is unique to differential forms. We shall associate with each differential form Φ of degree r, a differential (r+1) form, denoted by $d\Phi$, called the exterior derivative of Φ and $d\Phi \in T_{r+1}{}^*$. If Φ is of class C^k, then $d\Phi$ is of class C^{k-1}:

$$d\Phi = \sum_{1 \le i_1 \ldots i_r \le n+1} d(\Phi_{i_1 \ldots i_r})\, d\alpha^{i_1} \ldots d\alpha^{i_r} \tag{24}$$

For example, if Φ is a differential 1-form, $\Phi = \Sigma\, \Phi_i\, d\alpha^i$, its exterior differential is a 2-form:

$$d\Phi = \sum (\Phi_{ij} - \Phi_{ji})\, d\alpha_j \wedge d\alpha_i \tag{25}$$

If the differential form is of degree zero, for example, the function μ: N $\mapsto \Re$, then the differential of μ is the gradient of μ, which defines the vector field:

$$d\mu = \sum \left(\frac{\partial \mu}{\partial \alpha^i}\right) d\alpha^i \tag{26}$$

which corresponds to the 1-form Φ.

PROPERTIES OF EXTERIOR DERIVATIVES:

Let Φ be a differential form of class C^k ($k \ge 2$). Then

$$d\,(d\Phi) = 0 \tag{27}$$

That is, the operation d, applied twice, gives zero. We can see this easily for the simplest case of a differential 0-form. Let f be some differential 0-form; that is, f: N $\mapsto \Re$. Then

$$df = \sum \left(\frac{\partial f}{\partial \alpha^i}\right) d\alpha^i \tag{28}$$

from which we get

$$d\,(df) \;=\; \sum d\left\{\left(\frac{\partial f}{\partial \alpha^i}\right)\, d\alpha^i\right\} \tag{29}$$

Now $d\{(\partial f/\partial\alpha^i)\ d\alpha^j\} = \Sigma\ (\partial^2 f/\partial\alpha^j\ \partial\alpha^i)\ d\alpha^j$. Hence:

$$d\,(df) \;=\; \sum \left(\frac{\partial^2 f}{\partial \alpha^j\, \partial \alpha^i}\right) d\alpha^j\, d\alpha^i \tag{30}$$

The indices i,j vary independently in the summation. The terms with i = j are zero, since $d\alpha^i \wedge d\alpha^i = 0$. For $i \neq j$, group together the (i,j) terms and the (j,i) terms to obtain

$$d\,(df) \;=\; \sum \left(\frac{\partial^2 f}{\partial \alpha^i\, \partial \alpha^j}\right) d\alpha^i\, d\alpha^j \;+\; \left(\frac{\partial^2 f}{\partial \alpha^j\, \partial \alpha^i}\right) d\alpha^j\, d\alpha^i \tag{31}$$

From Young's theorem $\partial^2 f/\partial\alpha^i\ \partial\alpha^j = \partial^2 f/\partial\alpha^j\ \partial\alpha^i$, and we have seen that $d\alpha^i \wedge d\alpha^j = -\ d\alpha^j\ d\alpha^i$. Hence d (df) = 0. This result holds for differential forms of any degree.

Theorem of Frobenius:

Earlier, the integrability of a demand function, as given by the Frobenius' theorem, was defined in terms of a differential equation of the form

$$\frac{dm}{dp} \;=\; x(p,m) \tag{32}$$

Let x(p,m) be a given function of class C^1. A function $m = \mu(p)$ of class C^2 with values in $\Re$, such that

(i) $(p,\ \mu(p)) \in N$, for every p in Ω_n,

(ii) $\mu'(p) \;=\; x(p,\mu(p))$ for every $p \in \Omega_n$

is called a solution of equation (32).

Consider the differential 1-form Φ: $N \mapsto T^*(N,q)$. At the point $(p,m) \in \Omega$, and at the vector (dm, dp), it generates the vector $(dm - x(p,m)\ dp) \in E^{n+1}$. By this 'change of variable'

$$m = \mu(p) \tag{33}$$

it defines a differential form

$$\mu'(p) \cdot dp - x(p, \mu(p))\ dp \tag{34}$$

in N and with values in E^1. To say that $m = \mu(P)$ is a solution of the system of partial differential equations is to say that the differential form obtained by the change of variable (34) is zero.

This suggests a generalization of the notion of a solution of the partial differential equations in (32). This will be a system of two functions p(t), and m(t) of a variable t, such that the change of variable p = p(t), m = m(t) gives zero for the differential form. In other words,

$$\frac{dm}{dt} = x(p(t), m(t)) \cdot \frac{dp}{dt} \tag{35}$$

or the equality of two functions with values in the commodity space Ω^n, which is the tangent space T(N,q). In particular, if t is a real variable, we say that p = p(t), m = m(t) is an integral curve of equation (32) if these functions satisfy (35).

THE CLASSICAL FROBENIUS THEOREM: A STATEMENT

Consider the differential equation in (32): $dm/dp = x(p,m)$ where x is the given function of class C^1. We now require that for arbitrary $(p°,m°) \in \Omega$, this equation possess a solution $m = \mu(p)$ defined for p sufficiently close to p°, with the initial value $m° = \mu(p°)$.

THE FUNDAMENTAL THEOREM

For this to be the case, it is necessary and sufficient that for every point $(p,m) \in \Omega$,

$$X_p(p,m) + x_m(p,m) \cdot x(p,m)$$

define a symmetric element in T_2^*; that is,

$$\frac{\partial x_i}{\partial p_j} + \frac{\partial x_i}{\partial m} x_j = \frac{\partial x_j}{\partial p_i} + \frac{\partial x_j}{\partial m} x_i \tag{36}$$

for all p_i, p_j in Ω_n. When condition (36) is satisfied, we say that the equation (32) is completely integrable.

INTERPRETATION IN TERMS OF DIFFERENTIAL FORMS

Let us now define the problem in terms of differential forms. We have seen that the tangent hyperplanes to the manifold can be given locally by a differential 1-form Φ which is nowhere equal to the differential 0-form; that is $d\mu \neq 0$ on N. Its exterior differential is

$$d\Phi = -dx \wedge dp \tag{37}$$

where $d\Phi$ defines a bilinear mapping $T^* \times N \mapsto \Re$. Explicitly

$$d\Phi = -\left(\frac{\partial x}{\partial p}dp\right) \wedge dp - \left(\frac{\partial x}{\partial m}dm\right) \wedge dp \tag{38}$$

In this, replace dm by x(p,m) dp, to obtain the differential form of degree two:

$$\xi = -\left(\frac{\partial x_j}{\partial p_i} + \frac{\partial x_j}{\partial m} \circ x_i\right) + \left(\frac{\partial x_i}{\partial p_j} + \frac{\partial x_i}{\partial m} \circ x_j\right) \tag{39}$$

for any p_i p_j in Ω_n. The condition of complete integrability expresses the fact that this differential form ξ vanishes. That is, the hyperplanes given by dm/dp = x(p,m) is completely integrable, if the differential 2-form ξ, obtained by replacing dm by x(p,m) dp in $d\Phi$, is identically zero.

We shall now interpret this condition: take the distribution given as the tangent vector of $m = \mu(p)$. This tangent space T(N,q), which is the space of the commodity space, is of dimension n+1 with coordinates $\alpha^1 \ldots \alpha^{n+1} = p_1 \ldots p_n, m$. If the differential 1-form which is obtained from the dual $T^*(N,q)$ of the tangent space T(N,q), defined as the total differential of the function $m = \mu(p)$, is of dimension (n+1) - n = 1, and the field $T^*(N,q)$ defined in this way, is itself a distribution of dimension 1. That is, locally there exists a 1-form Φ, which generates $T^*(N,q)$.

THEOREM: Let Φ is a differential 1-form, defining a family of hypersurfaces such that $\Phi = \Sigma\, \Phi_i\, d\alpha^i$ whose coefficients do not all vanish simultaneously. In order that the equation $\Phi = 0$ be completely integrable, it is necessary and sufficient that it satisfy the integrability condition of $\Phi \wedge d\Phi = 0$ of Frobenius.

For $n = 3$, p_i, p_j, p_k,

$$x_i\left(\frac{\partial x_j}{\partial p_k} - \frac{\partial x_k}{\partial p_j}\right) + x_j\left(\frac{\partial x_i}{\partial p_k} - \frac{\partial x_k}{\partial p_i}\right) = x_k\left(\frac{\partial x_i}{\partial p_j} - \frac{\partial x_j}{\partial p_i}\right) = 0^{(40)}$$

where the left hand side of the equation is the coefficient of $dx_i \wedge dx_j \wedge dx_k$ in the canonical expression of the form $\Phi \wedge d\Phi$.

I.2

Surplus Measures As A Valid Index Of Utility Change

In this chapter, we address the question of the ability of a surplus measure to evaluate the effects of a change in the budget, that is consistent with the consumer's underlying preference ordering. That is, we examine the ability of the surplus measure to vary monotonically with changes in the utility function. Suppose a policy leads to changes from an initial point (p°,m°) to a new level (p',m'). If the two situations are such that the consumer prefers the new situation to the old, $(p',m')\ R^*\ (p^\circ,m^\circ)$ implying $V(p',m') \geq V(p^\circ,m^\circ)$, then the value of a surplus measure, evaluated from the old to the new situation for an appropriate path should be nondecreasing, in order that it provide a correct comparison between the two situations. In other words, we are looking at the question of the surplus measure providing a correct evaluation of a policy change purely in terms of a loss or gain; that is the sign of welfare change.

2.1 The Marshallian Surplus as a Valid Utility Index

The Marshallian surplus $\Phi(M)$ is measured as the change in utility valued at the marginal utility of income, at the point at which the change is taking place. For a given interval $I = [a,b]$, and a path $\alpha(t)$, $t \in [a,b]$, the Marshallian surplus $\Phi(M)$ is

$$\int \Phi(t) = \int \frac{1}{\lambda(t)}\, dV(t) \tag{1}$$

Valid Utility Index

The Marshallian surplus can correctly rank two situations in terms of a gain or a loss, if it has the same sign as the change in utility dV. The ability of $\Phi(M)$ to vary monotonically with dV depends on how λ varies along the path from the initial to the final point; that is on the sign of the derivatives of λ with respect to its arguments.

Consider the Lagrangian associated with the constrained utility maximization. The solution to this Lagrangian is a system of n+1 equations, which can be solved to obtain the demand function, and the Lagrangian multiplier λ, which is the marginal utility of income.

While the demand function is invariant under monotonic increasing transformation of U(x), the function $\lambda(p,m)$ is not. Hence the derivatives of λ with respect to its arguments are also not invariant for monotonic transformations of the utility function.

Let U be a utility function and λ^* the multiplier obtained from the constrained maximization of U. Consider a monotonic transformation of U, F[U(x)], where $F' > 0$. The corresponding transformation of λ^*, denoted by λ is

$$\lambda = \frac{\lambda^*}{F'} = \lambda^* \left(\frac{\partial U}{\partial F}\right) \tag{2}$$

Consider the derivative of λ with respect to p:

$$\frac{\partial\lambda}{\partial p} = \left(\frac{\partial U}{\partial F}\right)\left(\frac{\partial\lambda^*}{\partial p}\right) + \left(\frac{\partial^2 U}{\partial F^2}\right)\left(\frac{\partial\lambda^*}{\partial p}\right)^2 \quad (3)$$

While $\partial U/\partial F > 0$, and $(\partial\lambda/\partial p)^2 > 0$, there is no restriction on the sign of $\partial^2 U/\partial F^2$. Provided $\partial\lambda/\partial p \neq 0$, by a suitable transformation of the utility index, the derivative of λ with respect to p can be of any arbitrary magnitude and sign. In fact $\partial\lambda^*/\partial p$ need not be of the same sign as $\partial\lambda/\partial p$. Similarly $\partial\lambda/\partial m$ can be of any sign as well (Samuelson, 1947).

Hence, as utility varies along a path in the budget space, the marginal utility of income may either remain constant, or may vary along the path. For nonconstant λ, there is no a priori restriction that λ vary in any definite manner with dV.

CONSTANCY OF λ: If the marginal utility of income λ is constant over the path, then so is the 'converting factor' $1/\lambda$. This implies that not only does $\Phi(M)$ vary monotonically with dV, but the value of Φ is proportional to dV, providing a 'cardinal' measure of the surplus change. Thus, constancy of λ is a sufficient condition for the Marshallian surplus to be a valid index of surplus change. This is true for paths describing monotonic and nonmonotonic changes in utility. (The meaning of constancy of the marginal utility of income is discussed in chapter I.3.)

NON CONSTANT λ: Suppose λ is not constant, but varies along the path. Consider a strictly monotonic path of utility change, along which $dV(t)$ either monotonically increases or decreases. Since λ is strictly positive, $\Phi(t)$ will also monotonically increase or decrease along this path. Hence for strictly monotonic paths of utility changes, the Marshallian surplus will have the same sign as dV even if λ is not constant along the path.

However, for nonmonotonic paths in the budget space, if λ is not constant for this path, $\Phi(M)$ need not vary monotonically with dV. It is possible in such cases for the Marshallian surplus to give an incorrect ranking of the two situations. For example, for multiple price changes, where price increases are mixed with price decreases, the utility gain due to the price decreases may be greater than the loss in utility due to the price increase, giving a net welfare surplus. If λ is not constant with respect to these price changes, it may be possible that $\Phi(M) = 1/\lambda\ dV$ is negative, implying a welfare loss, when actually the consumer incurred a gain.

2.2 Invalidity Of The Compensating Variation For Nonmonotonic Paths

The Hicksian variations are defined as $\Phi(*) = de(P^*,U)$; $P^* = P'$ for the compensating variation, and $P^* = P^\circ$ for the equivalent variation. Since the function $e(P,U)$ is non decreasing in U, it follows that for monotonic changes in utility, $de(p^*,U)$ varies monotonically with dV, and both the Hicksian variations correctly compare the two situations, by providing the same sign as the utility change.

The situation however, is different for paths along which utility varies nonmonotonically. As before, consider the case of multiple price changes, where price increases are mixed with price decreases, and rectangular paths are used to describe these changes. For each segment of this path, along which utility varies monotonically, the Hicksian variations have the same sign as the utility change. However, for the entire path, the net effect would depend on the relative magnitudes of the decreses and increases in utility, as measured by the surplus measure.

Although both the Hicksian variations are derived from the expenditure function, the compensating vaiation allows only binary or pairwise comparisons. On the other hand, the equivalent variation can correctly compare several situations (McKenzie, 1983; Chipman and Moore, 1980; Hause 1975). This is because of the manner in which the two measures are constructed. The compensating variation measures the change in terms of the difference in the expenditure functions between the two utility levels, with the new price vector p' as the reference prices. Hence the surplus gain due to a price *decrease* from the initial situation will be measured using the *lower* prices as the reference prices. Similarly, the surplus loss due to price *increase* from the initial situation is measured using the *higher* prices as the reference prices. While each of these surplus changes are valid when the price change is compared to the initial situation, the magnitude of the surplus gain cannot be compared to that of the surplus loss.

As pointed out by Boadway and Bruce (1984), the only way in which this ambiguity can be avoided is if the surplus change due to the price increase is compared directly with that due to the price decrease, instead of comparing it to the base or initial situation. However this requires that the value of U° (or p°) in the expenditure function be appropriately modified, whenever a change in the direction of the price change takes place. This may not only be extremely cumbersome, but also not feasible, given the nature in which the Hicksian variations are measured, as discussed in subsequent chapters.

It has been shown (Chipman and Moore, 1980) that in some restrictive cases (homothetic preferences with constant income; and parallel preferences with constant price of the numeraire) the compensating variation can compare several situations. In the absence of these restrictions, the compensating variation cannot be used to measure surplus changes for paths along which utility varies nonmonotonically, unless the appropriate modifications are made to compare price increases directly with price decreases.

2.3 Equivalent Variation: A Valid Index Of Utility Change

In comparison, the equivalent variation uses the prices in the initial period as the reference level. Hence any number of situations can be compared with the base situation, and therefore with each other. It is therefore a valid numerical indicator of the consumer's preferences even for paths along which utility varies nonmonotonically.

2.3 A Numerical Illustration

We use a simple example to illustrate the possible ambiguities that may arise in using the Marshallian surplus and the compensating variation for nonmonotonic paths, when the underlying preferences are nonhomothetic. Consider the Linear Expenditure System (LES) obtained from the Stone Geary utility function, which is quasihomothetic (Stone, 1954; Deaton and Muellbauer, 1980). The corresponding demand functions are written as:

$$x_i = c_i - \left(\frac{\beta_i}{p_i}\right) \sum_k p_k c_k + \frac{\beta_i m}{p_i} \tag{4}$$

where $c_i \geq 0$, and $0 < \beta_i < 1$; $\Sigma\, \beta_i = 1$.

For the sake of illustration, consider two goods, x_1 and x_2. Let $\beta_1 = 0.25$, $\beta_2 = 0.75$; $c_1 = 12.75$, $c_2 = 6.4$. Further, let income m = \$100, and prices initially be at $p_1° = 1.0$, $p_2° = 1.0$.

Suppose the price of the first good p_1 increases from its initial value 1.0 to 1.8, and the price p_2 decrease from 1.0 to 0.7. The consumer surplus for this budget change, as measured by the equivalent variation:

$$EV = -[e(p^1,U^1) - e(p^0,U^1)] + m^1 - m^0 \qquad (5)$$

which for the Stone Geary system is:

$$\left(\frac{p_i^0}{p_i^1}\right)^{\beta_i} [\, m - \sum_j c_j\, p_j^1 \,] + \sum_j c_j\, p_j^0 - m \qquad (6)$$

is found to show a surplus gain of \$ 1.01839.

We calculated the Marshallian surplus, by integrating the demand function x_i with respect to p_i to get the area left of it between the price lines:

$$MS_i = -(\,\ln(p_i)\; c_j p_j\,) + c_i p_i)\,\beta_i + m\ln(p_i)\, c_i + c_i p_i \qquad (7)$$

We used a rectangular path, changing p_1 first and then p_2. The surplus loss due to the price increase is \$ - 21.40421, and the surplus gain due to the fall in price was found to be \$ 21.09135, showing a net surplus *loss* of \$ - 0.31286.

When the compensating variation was calculated,

$$CV = -[e(p^1,U^0) - e(p^0,U^0)] + m^1 - m^0 \qquad (8)$$

which for the Stone - Geary system is

$$\left(\frac{p_i^1}{p_0}\right)^{\beta_i} [\, m - \sum_j c_j p_j^0 \,] + \sum_j c_j p_j^1 - m \qquad (9)$$

showed a surplus *loss* of \$ - 0.90294.

For changes in the consumer's budget which result in non monotonic changes in utility, the Marshallian surplus and the compensating variation need not be a valid indicator of the surplus change, even in terms of measuring a surplus loss or gain. Unless preferences are homothetic, only the equivalent variation can correctly represent the welfare effects of the proposed change.

I.3

Surplus Measures And The Problem Of Path Dependence

We have seen that the changes in a consumer's utility in money terms, due to a budget change, can be measured as a line integral

$$\int_c \Phi(\cdot) \;=\; \int \Sigma\ \Phi_i(\alpha)\ d\alpha^i \tag{1}$$

The different surplus measures of Marshallian surplus and the Hicksian variations are obtained by taking the appropriate path α from the initial to final points. An important property of line integrals is that its value depends not only on the end points, but also on the path along which it is integrated. If the value of the integral varies from one path to another, then the line integral of Φ is said to be *path dependent*. However in some cases we may be able to find line integrals which are independent of the path of integration. In such cases the integral is path invariant, and its value is unique.

We have seen that for any surplus measure, several different paths may be defined for the interval I = [a,b]. Even for the subset of rectangular paths, it can be seen that different paths can be generated, by just changing the order in which the prices and income are changed. Since the choice of the order in which prices and income are changed is purely arbitrary, a surplus measure should give us the same value along any path; that is it should be path independent.

3.1 Poincaré's Theorem: Conditions For Path Independence

Let N (denoting the domain of the income compensation function μ) be a smooth manifold of dimension n+1. We have seen that the set of tangent vectors at some point q, denoted by T(N,q) is caled the tangent space of N at q. The space T_r^* of differential r-forms at q is the set of all r linear alternating functions Φ: $N \rightarrowtail T_r^*$. We have seen that a 0-form at q is, by convention, just a real number. The space T^* of 1-forms is the cotangent space to the point q at N. It is the space of linear functions on N, or the dual vector space to the tangent space T(N,q).

Consider the line integral in (1). We have seen that Φ, given as $\Sigma\ \Phi_i(\alpha)\ d\alpha^i$ is a smooth differential 1-form on the manifold N. From henceforth we shall refer to the mapping Φ as a differential 1-form, with all the related properties, as discussed in chapter I.1.

If $(\alpha^1 \ldots \alpha^{n+1})$ is the local coordinates of a neighborhood of q, then the tangent space T(N,q) has $\partial/\partial\alpha^1 \ldots \partial/\partial\alpha^{n+1}$ as its basis. The dual cotangent space has a dual basis, denoted by $(d\alpha^1 \ldots d\alpha^{n+1})$. The differential 1-form Φ thereby has the local coordinate expression: $\Phi = dm + \Phi_1\ dp_1 + \ldots + \Phi_n\ dp_n$, where each coefficient function Φ_i, i = 1,...,n has been shown to be the Marshallian demand function when Φ is the Marshallian surplus, or the appropriate Hicksian demand function $h(p,U^*)$, if Φ is the compensating or equivalent variation.

Let Φ be a non vanishing differential 1-form, defining a family of hyperplanes. Let Y denote a nonvanishing vector field. This field defines a tangent field ϕ, which at each point q, belongs to the hyperplanes of zeros of the 1-form $\Phi(q) = 0$. The field of zeros Φ will be invariant with respect to the phase flow of the field Y, if $\int \Phi$ is independent of the path of integration.

THEOREM 1: Let Φ be a differential 1-form of class C^1, that is continuously defined in a simply connected budget set, defining the change in utility in money terms. The integral of Φ, which defines a vector field in Ω, is independent of path, if Φ is the gradient of some scalar field f of class C^2 in Ω, called a *potential function* (Apostol, 1967).

It may be recalled that if Φ is a differential 1-form and of class C^1, then f is a differential 0-form of class C^2. A differential 0-form is none other than a scalar function in N.

If the budget set in which the line integral is defined is simply connected (see Introduction), then the condition in the above theorem is both necessary and sufficient for the integral to be path independent.

EXACT DIFFERENTIAL FORMS: We say that a differential 1-form Φ is *exact* in a simply connected space, if there is a function f of class C^2, such that $df = \Phi$ everywhere in Ω.

Consider a differential 1-form Φ: $N \mapsto T^*$. We now define conditions under which Φ is the gradient of a scalar field. That is, the conditions which guarantee the existence of a scalar function f

$$f : N \mapsto \Re; \quad df = \Phi \tag{2}$$

The necessary and sufficient conditions for (2) are given by:

POINCARÉ'S THEOREM: Let Ω be an n+1 dimensional manifold, and T^* the dual of the tangent space of a simply connected subset B of Ω. If $\Phi : N \mapsto T^*$, and satisfies

$$d\Phi = 0 \tag{3}$$

then there exists a scalar function, or a differential 0-form f of class C^2, where $f : B \in \Omega \mapsto \Re$, such that

$$df = \Phi \tag{4}$$

That is, if Φ is a differential 1-form that obeys $d\Phi = 0$, then it is the gradient field; and hence Φ is exact.

COROLLARY: If Φ is a differential 1-form which is exact, then it obeys $d\Phi = 0$. Hence the integral of Φ is independent of path in the simply connected subset G of Ω.

CONDITIONS FOR $d\Phi = 0$: Let Φ be a differential 1-form of class C^1, whose component functions Φ_i are nonvanishing. Suppose Φ satisfies $d\Phi = 0$. Then it follows that

$$d\Phi = \sum \left(\frac{\partial \Phi_i}{\partial \alpha^j} \right) d\alpha^j \wedge d\alpha^i = 0 \tag{5}$$

where $\wedge$ denotes exterior product (see chapter I.1). Note that the indices i,j vary independently in the summation. The terms with $i = j$ equal zero, since $d\alpha^i \wedge d\alpha^i = 0$. For the remaining terms, let us group together the (i,j) terms and the (j,i) terms to obtain

$$d\Phi = \sum \left(\left(\frac{\partial \Phi_i}{\partial \alpha^j}\right) d\alpha^j \wedge d\alpha^i + \left(\frac{\partial \Phi_j}{\partial \alpha^i}\right) d\alpha^i \wedge d\alpha^j \right) \tag{6}$$

Now, $d\alpha^j \wedge d\alpha^i = - d\alpha i \wedge d\alpha j$. Hence, we have

$$\frac{\partial \Phi_i}{\partial \alpha^j} = \frac{\partial \Phi_j}{\partial \alpha^i} \tag{7}$$

The corollary of this is that if a scalar function f exists, such that $df = \Phi$, then

$$d\Phi = d\,(df) \tag{8}$$

and hence

$$d\,(df) = \sum \left(\left(\frac{\partial^2 f}{\partial \alpha^i\, \partial \alpha^j}\right) d\alpha^j \wedge d\alpha^i = \left(\frac{\partial^2 f}{\partial \alpha^j\, \partial \alpha^i}\right) d\alpha^i\, d\alpha^j \right) \tag{9}$$

which from Young's theorem gives the equality of the cross partials, and hence $d(\,df\,) = d\Phi = 0$. Thus the condition for the path independence of the integral of Φ is the condition in (7) for $i,j = 1,\dots,n+1$.

IMPLICATIONS OF PATH INDEPENDENCE:

If Φ is a differential 1-form, then the following are equivalent:

(i) Φ is the gradient of some scalar function f in Ω,

(ii) If c_1 and c_2 are any two smooth curves in Ω, described by α^1 and α^2 respectively, with the same initial and final points a and b, then:

$$\int_{c_1} \Phi(\alpha^1) = \int_{c_2} \Phi(\alpha^2) \tag{10}$$

(iii) If c is a closed curve in Ω, then

$$\oint_c \Phi = 0 \tag{11}$$

3.2 Exact Forms And Integrating Factors

Consider a differential 1-form Ψ that is not exact; that is there is no scalar function, such that its gradient equals the differential form Ψ. However, it may be possible to find a multiplier ν, which if used to multiply the differential form Ψ, may result in a differential form that is exact:

$$d(\nu\,\Psi) = 0 \tag{12}$$

Now, it may be possible to find a scalar function F, such that

$$dF = \nu\,\Psi \tag{13}$$

The multiplier ν that converts the differential form Ψ into an exact equation is called an *integrating factor*.

EXISTENCE OF INTEGRATING FACTORS

A differential 1-form Ψ of class C^1, which is nowhere equal to the zero form on a manifold N, has coefficients Ψ_i, such that

$\Psi = \Sigma\Psi_i(\alpha)\ d\alpha^i$. For the form Ψ to possess an integrating factor ν, such that $\nu\Psi$ is exact it is necessary and sufficient that

$$\Psi \wedge d\Psi = 0 \tag{14}$$

THEOREM: Let Ψ be a form that satisfies conditions defined in (14), and ν be a function that vanishes nowhere. Then it implies that $d(\nu\Psi) = 0$; or that the form $\nu\Psi$ is exact.

PROOF: Consider the exterior derivative of the form $\nu\Psi$, that is $d(\nu\Psi)$. The exterior differential of a product is written as

$$d\ (\nu\,\Psi) = \nu \wedge (d\Psi) + (d\Psi) \wedge \Psi \tag{15}$$

Multiply (15) by Ψ. Now $\Psi \wedge \Psi = 0$. Hence,

$$d\ (\nu\,\Psi) \wedge \Psi = \nu \wedge (d\Psi) \wedge \Psi \tag{16}$$

Since exterior product is associative, it follows from (14) that $d\Psi \wedge \Psi = 0$. Hence, $d\ (\nu\Psi) \wedge \Psi = 0$. Since $\Psi > 0$, it follows

$$d\ (\nu\,\Psi) = 0 \tag{17}$$

and the conditions for path independence can now be written as:

$$\frac{\partial\,(\nu\,\Psi_i)}{\partial\alpha^j} = \frac{\partial\,(\nu\,\Psi_j)}{\partial\alpha^i} \tag{18}$$

3.3 Implications For A Path Independent Marshallian Surplus

Let Φ be a differential 1-form of class C^1. For a particular route of price - income changes, we can write Φ as

$$\Phi(M) = \frac{1}{\lambda}\,dV = dm - \sum_{i=1}^{n} x_i(p,m)\;dp_i \tag{19}$$

and the Marshallian surplus is obtained by integrating this differential form.

We have seen that the integrability conditions for the equation $\Phi = 0$ define the conditions for the existence of an indirect utility function V of class C^2, and a function λ of class C^1, where λ is the marginal utility of income, such that:

$$d\Phi = \lambda\,\Phi \equiv \lambda\,dm - \sum_{i=1}^{n} \lambda\,x_i(p,m)\;dp_i \tag{20}$$

Hence the equation $\lambda\,\Phi$ is exact, and implies:

$$d\,(\lambda\,\Phi) = 0 \tag{21}$$

Hence the path independence conditions in (18) imply

$$\begin{array}{llll} (a) & \dfrac{\partial(\lambda x_i)}{\partial p_j} = \dfrac{\partial(\lambda x_j)}{\partial p_i} & & \text{for } i,j = 1\ldots n \\ (b) & \dfrac{\partial(\lambda x_i)}{\partial m} = \dfrac{\partial\,\lambda}{\partial p_i} & & \text{for } i = 1\ldots n \end{array} \tag{22}$$

From this we get

$$\begin{array}{ll} (a) & \lambda\left(\dfrac{\partial x_i}{\partial p_j}\right) + x_i\left(\dfrac{\partial\lambda}{\partial p_j}\right) = \lambda\left(\dfrac{\partial x_j}{\partial p_i}\right) + x_j\left(\dfrac{\partial\lambda}{\partial p_i}\right) \\ (b) & \lambda\left(\dfrac{\partial x_i}{\partial m}\right) + x_i\left(\dfrac{\partial\lambda}{\partial m}\right) = \dfrac{\partial\lambda}{\partial m} \end{array} \tag{23}$$

If we impose the conditions that

$$
\begin{aligned}
&(a) \quad \left(\frac{\partial \lambda}{\partial p_j}\right)\frac{1}{x_j} = \left(\frac{\partial \lambda}{\partial p_i}\right)\frac{1}{x_i} \\
&(b) \quad \frac{\partial \lambda}{\partial m} = \left(\frac{\partial \lambda}{\partial p_i}\right)\frac{1}{x_i}
\end{aligned}
\tag{24}
$$

then we obtain

$$
\begin{aligned}
&(a) \quad \frac{\partial x_i}{\partial p_j} = \frac{\partial x_j}{\partial p_i} \\
&(b) \quad \frac{\partial x_i}{\partial m} = 0
\end{aligned}
\tag{25}
$$

a sufficient condition for the Marshallian surplus, given by the integral of (19) to be path independent.

The Marshallian surplus differs along different paths because the value of λ, and hence of the converting factor $1/\lambda$ varies along different paths. The imposition of the conditions in (25) implies that the variations of λ along different paths are the same, and hence the process of converting dV into money terms are identical along any path, giving a unique value of Φ.

A sufficient condition for (25) is that λ be constant with respect to the variables that change. That is

$$
\begin{aligned}
&(a) \quad \frac{\partial \lambda}{\partial p_i} = 0 \quad \text{for } i = 1....n \\
&(b) \quad \frac{\partial \lambda}{\partial m} = 0
\end{aligned}
\tag{26}
$$

It is well known that since it is homogeneous of degree minus one, λ cannot be constant with respect to all its arguments (Samuelson 1942). Conditions in (26) cannot therefore hold simultaneously; constancy of λ can have meaning only if the price income changes are restricted to some subset of the n+1 price - income space. The following possibilities have been considered (Silberberg 1972; Burns 1972; Chipman & Moore 1980).

(1) All n prices change, but income remains constant. The path independent conditions relevant are 26 (a). In this case λ is a function only of M, then λ is 'constant' with respect to the variables that change. It has been shown by Samuelson that λ is a function of M and independent of all prices only if the underlying preferences are homothetic.

(2) Only n-1 prices and income change; and one price (say n), remains unchanged. Then conditions 26 (a) for 1...n - 1 and 26 (b) hold. If λ is a function only of p_n, it is constant with respect to the variables that change. It has been shown that this implies that preferences are quasilinear with respect to the nth good.

(3) Some subset of prices $p_1 \ldots p_k$ change, and $p_{k+1} \ldots p_n$ prices remain unchanged. In this case conditions 26 (a) hold only for i,j = 1...k. This requires that λ not be a function of the $p_1 \ldots p_k$ prices and hence is constant with respect to the changing variables.

It should be emphasized that when we require some variables to remain constant, we imply that they remain unchanged for the entire path.

The path independence conditions imply that that the process of converting dV into money units is identical for all paths, giving a unique value of Φ. For example, let prices p_i and p_j change. Since

$$\frac{\partial \lambda}{\partial p_i} = \lambda \left(\frac{\partial x_i}{\partial m}\right) + x_i \left(\frac{\partial \lambda}{\partial m}\right) \tag{27}$$

it follows from (23) that

$$\frac{\partial \lambda}{\partial p_i}\frac{1}{x_i} = \frac{\partial \lambda}{\partial p_j}\frac{1}{x_j} \Rightarrow \left(\frac{\partial x_i}{\partial m}\right)\frac{1}{x_i} = \left(\frac{\partial x_j}{\partial m}\right)\frac{1}{x_j} \tag{28}$$

from which it follows that the income elasticities of the two goods are equal. However, if λ is not constant, the converting factor $1/\lambda$ depends on the path of integration, and the value of $\Phi(M)$ would differ along different paths. The magnitude of the divergence in the values of the Marshallian surplus depends upon the difference in the values of the income elasticities for the goods in the region of the budget space considered.

3.4 The Path Independent Hicksian Variations

The Hicksian variations in terms of the differential form Φ can be written as

$$\Phi(*) = dm - \sum_{i=1}^{n} h_i(p, U^*)\, dp_i \qquad U^* = U^0 \text{ for } CV; \quad U^* = U^1 \text{ for } EV \tag{29}$$

giving the compensating and equivalent variations respectively. Since $\Phi(*) = 1/d\lambda\ dV = de(P,U^*)$, the corresponding differential forms are gradients of a scalar function. The equation $\Phi(*)$ in

(29) is exact and hence the path independece conditions

$$\frac{\partial h_i}{\partial p_j} = \frac{\partial h_j}{\partial p_i} \tag{30}$$

are always satisfied. This is easily seen from the fact that the existence of the expenditure functions $e(P,U^*)$ implies that

$$\begin{aligned} d\Phi(*) &= d\,(de) \\ \therefore \quad \frac{\partial h_i}{\partial p_j} &= \frac{\partial^2 e}{\partial p_i\,\partial p_j} \end{aligned} \tag{31}$$

Hence from Young's theorem (30) follows. The compensating and equivalent variations are independent of path and are therefore uniquely defined. Therefore they can be measured along any arbitrary path in Ω for the interval $I = [a,b]$.

Part III

ON THE MEASUREMENT PROBLEM

II.1

Approximating The Equivalent Variation

We have seen that the Marshallian surplus need not be a valid measure of surplus change, for budget changes that lead to nonmonotonic changes in utility. It is a valid measure only if we impose restrictions either on the underlying preferences, or on the of paths of integration. In the absence of these restrictions, the Marshallian surplus does not bear a one-to-one correspondence with changes in the utility, and hence fails to provide a correct ranking of different situations.

On the other hand, the equivalent variation can provide a valid ranking of different situations that are consistent with the consumer's underlying preferences. However, unlike the Marshallian surplus which can be computed from ordinary demand functions, the equivalent variation requires knowledge of the expenditure function, or the Hicksian demand functions, which are not directly observable.

Two developments in the literature have evolved to cope with this problem. On the one hand, practitioners of cost benefit analysis have continued to use the Marshallian surplus, on the grounds that it provides a good approximation to the Hicksian measures. Such an approach has been justified by the work of Willig (1976), who has derived precise upper and lower bounds on the percentage errors incurred in approximating the Hicksian variations with the Marshallian surplus. He has shown that for a wide range of circumstances, the Marshallian surplus varies only slightly from path to path, and for any path, will closely approximate the Hicksian variations. The other and more recent approach has been to find a methodology to directly calculate the Hicksian variations using only Marshallian demand functions (McKenzie 1976; Hausman 1981; Vartia 1983), which is discussed in subsequent chapters.

In this chapter we consider the approximation problem in the context of budget changes that lead to nonmonotonic changes in utility. It is our aim to point out, that the approximation results of Willig hold only in those situations that lead to monotonic changes in utility. The cases considered by Willig are single price changes, and multiple price changes where all prices move in the same direction. Since Willig considers only rectangular paths, these budget changes clearly lead to changes in utility that are monotonic. As cautioned by Willig himself, when price increases are mixed with price decreases, the "..compensated demands can stray significantly from Marshallian demands along the path of integration and consequently the Marshallian surplus can fail to closely approximate [sic] the compensating and equivalent variations" (1979, pg. 472).

1.1 The Approximation Results For Monotonic Paths

Consider the Marshallian surplus and the equivalent variation. The difference in the two measures lies in the fact that the Marshallian surplus is measured as the area left of the ordinary demand functions between the price lines, while the equivalent variation is the corresponding area left of the compensated demand functions. As seen in the Slutsky equation, the Hicksian demand functions diverge from the Marshallian ones by the compensation in income that is continously made for the former as prices change, while income is kept unchanged for the latter. The difference in the two surplus measures depends on this income effect; if there is no income effect, it follows that the Marshallian measure equals the Hicksian variation.

The divergence of the two measures can be quantified in terms of two factors: the proportion of the consumer's income spent on the goods whose prices change, as well as the size of their income elasticities. This was first pointed out by Hotelling, and was later given a precise formulation by Willig, in terms of upper and lower bounds on the percentage deviation of the Marshallian surplus from the Hicksian variations.

Let x denote the vector of commodities, and $\epsilon_i(p,m)$ the income elasticity of the ith good. Consider a multiple price change, where the prices are increased from some initial value p_i° to a final value p_i', for $i = 1,....n$, and income remains unchanged at $m^\circ = m' = m$. We shall consider only rectangular paths; hence such paths describe strictly monotonic changes in utility.

Although several rectangular paths are possible, Willig considers a specific one along which the surplus measures are calculated: the n prices are changed in order from their initial to their final values, one at a time, while those that have moved, stay at their new values. Consider a linear segment α_i of this path, where the price of the ith good (p_i) is being changed, all other prices remain unchanged either at their new or initial values:

$$\alpha_i = \{ p: (p_1^1, \ldots, p_{i-1}^1, p_i, p_{i+1}^0, \ldots, p_n^0); \quad p_i = (1 - t)p_i^1 + tp_i^0\} \tag{1}$$

For this segment of the path, the equivalent variation is:

$$EV_i = \mu(p_1^1, \ldots, p_{i-1}^1, p_i^1, p_{i+1}^0, \ldots, p_n^0; p^1 m^1) - \mu(p_1^1, \ldots, p_{i-1}^1, p_i^0, p_{i+1}^0, \ldots, p_n^0; p^1 m^1) \tag{2}$$

The Marshallian surplus (MS_i) is measured as

$$MS_i = \int_{p_i^0}^{p_i^1} x_i(p_1^1 \ldots p_{i-1}^1, p_i, p_{i+1}^0, \ldots, p_n^0; m)\, dp_i \tag{3}$$

(Note that for a price increase the EV_i would have a negative sign). The equivalent variation for the entire path is the sum of the EV_i for each segment of the rectangular path, where the appropriate compensation in income is made for each price change. Similarly, the total Marshallian surplus is the sum of the changes for each segment $MS = \Sigma MS_i$; that is, it is the sum of the areas left of the ordinary demand curves between the price lines, sequentially shifted as prices change, one at a time, from their initial to their final values.

Approximations

To begin with, the relationship between the Marshallian surplus and the equivalent variation is derived. From this, the bounds for the percentage error in approximating the equivalent variation with the Marshallian surplus is obtained.

Consider the segment α^i of the path where only the price p_i is changed. The income elasticity of the good x_i,

$$\epsilon_i = \frac{\partial x_i}{\partial m} \cdot \frac{m}{x_i} \tag{4}$$

can be rewritten as:

$$\frac{1}{x_i} dx_i = \frac{\epsilon_i}{m} dm \tag{5}$$

Since we are considering changes only in the price p_i, let us simplify the notation as follows:

$$p_i^1 = p_1^1, \ldots, p_{i-1}^1, p_i^1, p_{i+1}^0, \ldots, p_n^0$$

and

$$p_i^0 = p_1^1, \ldots, p_{i-1}^1, p_i^0, p_{i+1}^0, \ldots, p_n^0$$

For simplicity of notation let $m^1 = \mu(p^1;p^1,m)$, $m^2 = \mu(p^\circ;p^1,m)$. Since income does not change, $m^1 = m^0 = m$. Integrate (5) with respect to m^1 and m^2,

$$x_i(P,m^2) = x_i(p,m^1) \exp \int_{m^1}^{m^2} \frac{\epsilon_i}{m} dm \tag{8}$$

Approximations

Consider the income compensation function $\mu(p;\ p,m)$. The derivative of μ with respect to price p_i gives the following differential equation

$$\frac{\partial l(p^0;\, p^1,m)}{\partial p_i} \;=\; x_i(p,\ \mu(p^0;\, p^1,m)) \;\equiv\; x_i(p,m^2) \qquad \textbf{(9)}$$

for initial conditions $\mu(p°;\ p°,m) = m^0 = m$. Substituting (8) into (9), and given that $m^2 = \mu(p°;p^1,m)$ and $m^1 = \mu(p^1;\ p^1,m)$,

$$d\mu(p;\, p^1,m^0) \;=\; x_i(p,m)\ dp_i\ \ \exp \int_{m^1}^{m^2} \frac{\epsilon_i}{m}\ dm \qquad \textbf{(10)}$$

The equation (10) can be used to find the relationship between the EV_i and the MS_i. For example, if ϵ_i equals zero, then EV_i defined on the left hand side equals the MS_i. If preferences are homothetic, such that ϵ_i equals one, then

$$\mu(p^0;\, p^1,m) \;=\; m\ \ \exp\left[\ \frac{1}{m}\int_{p_i^1}^{p_i^0} x_i(p,m)\ \ dp_i\ \right] \qquad \textbf{(11)}$$

from which the EV_i in terms of the MS_i can be calculated.

For the general case of nonhomothetic preferences: let ϵ_i be constant in the price income region under consideration. Then equation (10) reduces to:

$$\mu(p^0;\, p^1,m) \;=\; m\left[\ 1 + \frac{1-\epsilon}{m}\int_{p_i^1}^{p_i^0} x_i(p,m)\ \ dp_i\ \right]^{\frac{1}{1-\epsilon}} \qquad \textbf{(12)}$$

and the following relationship between the equivalent variation and the Marshallian surplus follows:

$$\frac{MS_i - EV_i}{MS_i} \approx \frac{\epsilon_i \, MS_i}{2m} \tag{13}$$

However, if ϵ_i is not constant along the path, an explicit solution for (10) is not possible. In this case, ϵ_i^1 and ϵ_i^2, denoting the lower and upper values of the income elasticities of good x_i, can be used to derive bounds on the divergence of the Marshallian surplus from the Hicksian variations. Willig shows that for an increase in the price p_i, the bounds on the deviation of MS_i from the EV_i is defined by

$$\frac{\epsilon_i^1 \, |MS_i|}{2m} \le \frac{MS_i - EV_i}{|MS_i|} \le \frac{\epsilon_i^2 \, |MS_i|}{2m} \tag{14}$$

If all prices are increased, then for each linear segment of the path α_j, j = 1...n, the above inequality holds, where the bounds are determined by the elasticities of the good whose price is being changed for that segment. Given that $\mu(p;\ p'm)$ is a monotonic function of p;

$$\mu(p_1^1 \ldots p_k^1,\ p_{k+1}^0, \ldots p_n^0) < \mu(p_1^1 \ldots p_{k-1}^1,\ p_k^0 \ldots p_n^0) \tag{15}$$

and $\mu(p';\ p'm) = m$; these inequalities can be summed to yield:

$$\frac{\epsilon_* \, |MS|}{2m} \le \frac{MS - EV}{|MS|} \le \frac{\epsilon^* |MS|}{2m} \tag{16}$$

where $\epsilon_* = \min\ \{\epsilon_i^1\}$ and $\epsilon^* = \max\ \{\epsilon_i^2\}$ are bounds on the income elasticities of all the goods whose prices change and $MS = \Sigma MS_i$.

If all prices decrease, then the EV_i exceeds the MS_i, and the bounds on the percentage deviation of the EV from the MS; that is (EV - MS) / MS are identical to (16).

Willig has shown that the percentage deviations of the Marshallian surplus from the equivalent variation will be small if the MS_is are a small fraction of the consumer's income, and if the income elasticities ϵ_i are not large.

1.2 Approximation Results For Nonmonotonic Paths

Consider the case of multiple price changes, where price increases are mixed with price decreases. Although the bounds on the percentage deviation between the Marshallian surplus and the equivalent variation in (14) with the appropriate signs, apply to each segment α_i of the path, these inequalities can no longer be summed up to provide bounds for the deviation of the two measures for the *entire* path. This is because the changes in prices along such a path no longer lead to monotonic changes in the income compensation function. Hence we cannot infer about the closeness of the two measures merely from the size of MS/m and the income elasticities of the goods. It is possible for the Marshallian surplus to deviate substantially from the equivalent variation even if MS/m and the income elasticities were small. A numerical example is provided at the end of this chapter.

A nonmonotonic path can be viewed as consisting of two segments, along each of which, changes in the budget lead to monotonic changes in the income compensation function. For multiple price changes, where price increases are mixed with price decreases, this can be done by partitioning the vector of prices into two subsets such that one contains all prices whose value increases, while the other groups all those whose value decreases. Then each of these paths, describing either price increases or decreases, are ones along which utility, and hence the function $\mu(\cdot)$, varies monotonically. The approximation bounds in (14) can now be applied for each of these sub paths.

Let $p_1 \ldots . p_k$ be the prices whose values increase, and $p_{k+1} \ldots . p_n$ be those whose value decreases. Let MS^I denote the Marshallian surplus for the $p_1 \ldots p_k$ price increases and MS^D for the $p_{k+1} \ldots p_n$ price decreases. Using the notations in (6) and (7), let $\mu(p_n) \equiv \mu(p_i^{\,1}) - \mu(p_i^{\,0})$ denote the compensated change in income associated with the change in the price p_i. Then the EV for the entire price change is

$$\mu(p^1) - \mu(p^0) \equiv \mu(p_n^{\,1}) - \mu(p_1^{\,0}).$$

The relationship between the Marshallian surplus and the equivalent variation can be broken into:

$$\begin{gathered} \mu(p^1) - \mu(p^0) - (MS^D - MS^I) = \\ (MS^I - (\mu(p_k) - \mu(p^1)) + \mu(p_k) - \mu(p^0) - MS^D) \end{gathered} \tag{17}$$

First consider the path along which prices increase:

$$p_k^1 > p_k^0, \ldots, p_1^1 > p_1^0 \Rightarrow \mu(p_k) < \ldots . < \mu(P^0) \tag{18}$$

For this segment of the path, the Marshallian surplus exceeds the equivalent variation: $\mu(p_k) - \mu(p^\circ) < MS^D$; and the percentage deviation between the two measures are as in (16), with the most extreme of the income elasticities of the k goods providing the bounds.

Consider now the remaining path along which the n - k+1 prices decrease. Again the path describes changes along which the income compensation function varies monotonically:

$$p_n^1 < p_n^0, \ldots, p_{k+1}^1 < p_{k+1}^0 \rightarrow \mu(p_n) > \ldots > \mu(p_k) \qquad \textbf{(19)}$$

and the equivalent variation for these price changes is $\mu(p_n) - \mu(p_k)$. However, the bounds in (16) no longer apply for these price changes. This is because the equivalent variation for this path is measured using $\mu(p_k)$ as the base income, whereas the income (m) is used to calculate the bounds in (16). Since this segment of the path is preceded with price increases, the compensated income $\mu(p_k)$ is smaller than m. If the percentage deviations of the two measures is calculated using m and not $\mu(p_k)$, it will understate the difference in the two measures. However $\mu(p_k)$ is not known, and unless appropriate adjustments for the income (m) are made so that it approximates $\mu(p_k)$, the bounds as derived in (16) can no longer apply for this segment of the path.

Now consider the Marshallian surplus, which for the entire path consisting of all price changes, is equal to the sum of MS^I

and MS^D. Since these two components are of opposite signs, it may be possible for the net MS to be small, even though the components MS^I and MS^D are not. However, the smallness of the deviation of the two measures for the entire path depends not on the net MS, but on the absolute value of MS^I and MS^D. The price changes as a whole may have a moderate effect on the welfare of the consumer, but if the price increases or price decreases by themselves result in large income effects, the Marshallian surplus will fail to provide a good approximation to the equivalent variation. Hence for price changes that lead to nonmonotonic changes in the income compensation function, the smallness of the net MS/m, or of the income elasticities is not sufficient to guarantee that the two measures will be close in value.

1.3 A Numerical Example

Consider again the demand functions corresponding to the Linear Expenditure System obtained from the Stone Geary utility function:

$$x_i = c_i - \left(\frac{\beta_i}{p_i}\right) \sum_k c_k p_k + \frac{\beta_i m}{p_i} \tag{20}$$

For the sake of illustration, consider two goods x_1 and x_2, whose marginal budget shares are $\beta_1 = 0.22$ and $\beta_2 = 0.78$, and income m = 100. Let $c_1 = 14.6$ and $c_2 = 3.2$, and the prices initially be at $p_1° = 1.0$ and $p_2° = 1.0$.

Approximations

Suppose the price of both goods increase, say $p_1' = 1.25$ and $p_2' = 1.3$. For these ranges of prices, the extremes of the product specific income elasticities are $\epsilon_* = \text{Min}\ \{0.62,\ 1.16\}$ and $\epsilon^* = \text{Max}\ \{0.67,\ 1.21\}$.

The Marshallian surplus calculated sequentially for these changes in p_1 and p_2 showed a loss of 7.6 and 16.94, giving a total of 24.54. The equivalent variation was found to be 22.0. Hence (EV - MS) / MS is (24.54 - 22.0) / 24.54 = 0.1035. The bounds on the percentage deviation of the MS from the EV shows a range of 0.08 and 0.11, which is consistent with the results.

Consider now the case of price increases mixed with price decreases. Let $p_1' = 1.5$ and $p_2' = 0.7$. For these prices, the income elasticities range between 0.6 and 1.3. Let us first consider the price increase. The MS^I showed a surplus loss of 14.32878. For the price decrease, the MS^D showed a surplus gain of 21.23513. Hence the net MS for the entire price change is a gain of 7.61035.

If we calculate the bounds on the percentage deviation of this net MS from the EV using income elasticities, it shows a very small deviation in the two measures, ranging between 0.026 and 0.045. However, when the EV was actually calculated for these price changes, it showed a surplus gain of 9.44. Hence the percentage error due to the approximation of the equivalent variation by the Marshallian surplus is actually quite large;

$$(EV - MS) / MS = (9.44 - 7.61) / 7.6 = 24.1$$

Thus the value of the net MS does not give any indication of the size of the errors when price increases are mixed with price decreases.

Consider the equivalent variation for each segment of the path. For the price increase the EV was a loss of 13.69, and for the price fall, a gain of 23.13. The difference between the two measures for the entire path can be obtained from the components:

$$EV - MS = 14.33 - 13.69 + 23.13 - 21.94 = 1.83.$$

Hence the percentage deviation of the EV from the MS is 24.1.

II.2

Measuring The Equivalent Variation Directly: Linear Expansion Paths and Integrating Factors

---

We have seen that for paths in budget space along which utility changes nonmonotonically, deviations of the Marshallian surplus from the equivalent variation may be substantial and may even provide an incorrect ranking of the situations. Hence the use of the Marshallian surplus to measure welfare changes in such situations may be unjustifiable. It is a common belief that the computation of the Hicksian measures is intrinsically more difficult since it requires knowledge of the compensated demand curves: "....it is usually rather difficult to calculate the compensated demands, for they require that our data on price and income be integrated twice, which in turn requires not only integrability but a large amount of information not commonly available, whose normal content of errors will necessarily be increased by these complicated manipulations" (Seade 1978 p 511). The aim here is to calculate the equivalent variation directly from the observable demand functions. In this chapter, we restrict attention to demand functions whose expansion paths are linear. The class of preferences giving rise to such paths has been extensively studied by Afriat (1972), and will be the basis of the analysis here.

2.1 Implicit Definition Of The Equivalent Variation

The equivalent variation is defined as the amount of income required instead of the price change being contemplated, which would produce the same utility change as the latter:

$$V(p^1,m^1) = V(p^0,m^0 + EV) \tag{1}$$

which can be rewritten as

$$V(p^1,m^1) - V(p^0,m^1) = V(p^0,m^0 + EV) - V(p^0,m^1) \tag{2}$$

The two sides give the same change in utility brought about by the change in prices and the compensated income respectively. Using Roy's identity, and the relation $\lambda = \partial V / \partial M$, we obtain:

$$\int_{m^0+EV}^{m^1} \lambda(p^0,m)\ dm = \int_{p^0}^{p^1} \sum_{i=1}^{n} \lambda(p,m^1)\ x_i(p,m^1)\ dp_i \tag{3}$$

which defines the equivalent variation implicitly in terms of the ordinary demand functions.

Solving the equation in (3) for the equivalent variation requires knowledge of the marginal utility of income $\lambda(p,m)$. The integrability conditions only guarantee the existence of the integrating factor; neither the functional form of λ, nor its values over the budget changes are known or observed. Hence to solve for the equivalent variation from (3) requires that the function $\lambda(p,m)$ be calculated from observable data.

The function $\lambda(p,m)$, defined as $\partial V(p,m)/\partial m$ is a mapping in the n+1 dimensional budget space:

$$\lambda : \ \Omega_n \, x \, E^1 \ \mapsto \ \Omega \tag{4}$$

The derivative of λ with respect to the price vector defines the following n partial differential equations:

$$\begin{aligned} \frac{\partial\lambda}{\partial p_i} &= \frac{\partial \ (\partial V / \partial m)}{\partial p_i} \\ &= - x_i \frac{\partial\lambda}{\partial m} - \lambda \frac{\partial x_i}{\partial m} \end{aligned} \tag{5}$$

However to solve for the λ as a function of n+1 variables, the above system is indeterminate, unless one more equation can be obtained. One possibility is to place restrictions on λ such that its movement in the budget space is constrained, making it independent of one or more variables. The preferences that result in such restrictions in λ are those which give rise to linear expenditure functions and straight line expansion paths.

2.2 Linear Expenditure Functions And Expansion Paths

Consider the vector of commodities x defined in Ω^n, and the vector of prices p in Ω_n. Let us define $y = p^T x$. The product $p^T x$ defines an n dimensional space; $y: \Omega_n \times \Omega^n \mapsto E^n$. Let S denote a subset of the commodity space Ω^n, such that for all $x \in S$, $U(x) \geq U°$, for some level of utility $U°$. Then the set $y(S)$ corresponding to $x \in S$, defining the expenditure: $X \in S$, $P > 0$:

$$y \ = \ [\ p^T x : \ p^T x = y(x) , \ \ x \in S \ ] \tag{6}$$

is the set of values attained by y in $S \in \Omega^n$. The set $y(S)$ is nonempty; $S > 0$, $p > 0$. Further $p^T x$, and hence the set $y(S)$ is bounded from below, which implies that the epigraph of this set has a lower limit. The greatest lower bound of y(S) defines the expenditure function $e(p, U^*)$ as the lower limit of expenditure at prices p of attaining the utility of a bundle x. It implies that the cost $p^T x$ of x at the prices p is also the minimum cost of attaining the utility of x.

Consider the function $y = p^T x$. It is positively homogeneous of degree one in p; ($\iota y = (\iota p)^T x$, $\iota > 0$). Hence the epigraph of the function y is a cone. Thus for every $p \in \Omega_n$, the half line D_p, given by

$$D_p = \{ \iota p, \ \iota > 0; \ p \in \Omega_n \} \tag{7}$$

is contained in e(p,U). Hence e(p,U) is a union of half lines coming through the origin for $p > 0$. Therefore the expenditure function is linearly homogeneous and concave in p.

Let R be a preference ordering for which the associated expenditure function which takes the form:

$$e(p,U) = a(p)\ \theta(x) + b(p) \tag{8}$$

called a *linear* expenditure function (Afriat, 1972). The function e(p,U) is increasing in U for every p in Ω_n. Hence for any p, $\theta(x)$ is a utility function which represents R; $x^1 R\, x^2 \equiv \theta(x^1) \geq \theta(x^2)$. Since the order R is concave, the function θ representing it is quasiconcave.

Consider the indirect utility function associated with $\theta(x)$, obtained by inverting the expenditure function $e(p,U) = p^T x \equiv m$

$$V(p,m) \quad = \quad \frac{m - b(p)}{a(p)} \tag{9}$$

The marginal utility of money, defined as the partial derivative of the function V with respect to m:

$$\lambda \quad = \quad \frac{1}{a(p)} \tag{10}$$

is independent of income: $\partial\lambda/\partial m = 0$. Hence, linearity of the expenditure function implies that they admit representation by a utility function, such that the corresponding partial derivative with respect to income, that is the marginal utility of income, is independent of income.

The converse is also true. Consider the inverse of λ: the marginal cost of utility:

$$\frac{\partial e(p,U)}{\partial U} \quad = \quad \frac{1}{V_m} \quad \equiv \quad a(p) \tag{11}$$

which upon integrating yields $e(p,U) = a(p)U + c$ where $c = b(p)$ is the constant of integration.

The partial derivative of V with respect to m is linear, and is therefore concave. Hence the quasiconcavity of $\theta(x)$ can be replaced by strict concavity.

The term b(p) represents overhead or fixed costs, giving the expenditure when utility $\theta(x)$ is zero. Since e(p,U) is linearly homogeneous in p, both a(p) and b(p) are also linearly homogeneous.

Consider the expansion locus:

$$E_p = [\, x \;:\; x E_p \,] \qquad \textbf{(12)}$$

defined as the set of x, which for a given preference ordering R and prices p, is the utility maximizing quantities demanded as income m changes.

The demand function x(p,m), derived from the constrained maximization of the utility function, is therefore the point of intersection of the expansion set E_p with the indifference surface. Since the utility function is strictly concave, the corresponding indifference surfaces are convex. Hence there is a unique x for every level of income and the expansion loci are therefore paths.

Consider two points x^1Ep and x^2Ep in the expansion set Ep. By definition, x^1 and x^2 are utility maximizing bundles when prices are p for different levels of income. By the assumption of strict concavity of the preference ordering R, these utility maximizing bundles are unique at these levels of income. Hence:

$$\begin{array}{ll} (i) & a(p)\ \theta(x^1) + b(p) = p^T x^1 \\ (ii) & a(p)\ \theta(x^2) + b(p) = p^T x^2 \end{array} \qquad \textbf{(13)}$$

Consider $\alpha > 0$, $\beta > 0$, such that $\alpha + \beta = 1$. Let x denote the line segment joining x^1 and x^2; $x = \alpha x^1 + \beta x^2$. Therefore $\theta(\alpha x^1 + \beta x^2) = \theta(x)$. By the strict concavity of θ, we have

$$\theta(x) \geq \alpha\, \theta(x^1) + \beta\, \theta(x^2) \qquad \textbf{(14)}$$

Consider the expenditure functions in (13). Multiply (13) (i) by α and (ii) by β, and the sum of the two gives

$$a(p)[\alpha\,\theta(x^1) + \beta\,\theta(x^2)] + b(p)[\alpha + \beta] = p^T(\alpha x^1 + \beta x^2) \quad (15)$$

Hence

$$a(p)[\alpha\,\theta(x^1) + \beta\,\theta(x^2)] + b(p) = p^T x \quad (16)$$

But given the inequality in (14) we can obtain

$$a(p)\,\theta(x) + b(p) \geq p^T x \quad (17)$$

In general, for any bundle $x \in \Omega^n$, $e(p,U(x) \leq p^T x$, implying

$$a(p)\,\theta(x) + b(p) \leq p^T x \quad (18)$$

From (17) and (18), it follows that

$$a(p)\,\theta(x) + b(p) = p^T x \quad (19)$$

Hence the commodity bundle x is also a utility maximizing bundle at prices p, implying xEp, or that x is also a point in the expansion set Ep. Since the line segment joining the points x', $x^2 \in$ Ep is contained in Ep, therefore it follows that the set Ep is convex.

When the preference ordering R is strictly concave, the expansion sets are paths; which implies that it intersects the indifference surfaces at a unique point. Hence for the expansion set Ep to be convex, it must be a straight line.

Consider the converse of this: let x' and x^2 be two points in an expansion set. If the expansion set between these two points is convex, then the corresponding expenditure function is linear.

For some prices p, let Ep denote the expansion set, such that x^1, $x^2 \in$ Ep. By definition $x^1$and x^2 are utility maximizing bundles at prices p. Let us denote the utility associated with x^1 and x^2 by U^1 and U^2 respectively.

Let x_t be a commodity bundle obtained by linearly combining x^1 and x^2

$$\begin{aligned} x_t(p) &= (1-t)x^1(p) + t x^2(p) \\ &= x^1(p) + (x^2(p) - x^1(p))\, t \end{aligned} \quad \textbf{(20)}$$

at prices p, where $0 \leq t \leq 1$. Given that the set Ep is convex, $x_t(p)$ is also contained in Ep. Since $x^1(p)$ is a utility maximizing bundle at prices p, it implies $x^1(p)$ R $x^1(p^*)$ for all $p^* \in \Omega_n$. Hence

$$p^T x^1(p) = Min\,[p^T x : U(x) \geq U(x^1(p))] \quad \textbf{(21)}$$

and

$$p^T x^1(p) \leq p^T x^1(p^*) \quad \textbf{(22)}$$

Similarly, for the utility maximizing bundle $x^2(p)$, we have

$$p^T x^2(p) \leq p^T x^2(p^*) \quad \textbf{(23)}$$

Multiply (22) by t and (23) by (1 - t), and add to obtain

$$p^T x_t(p) \leq p^T x_t(p^*) \quad \textbf{(24)}$$

$0 \leq t \leq 1$. By the convexity of Ep, $x_t(p) \in$ Ep. Therefore

$$x_t(p) \;\; R \;\; x_t(p^*) \quad \textbf{(25)}$$

which implies that the bundle $x_t(p)$ is a utility maximizing bundle, and is the point of intersection between Ep and the indifference surface corresponding to the utility level U_t. It therefore follows that $\theta(x_t(p)) = t$, and $\theta(x)$ is a utility function representing the preference ordering R.

Consider the total expenditure associated with utility t at prices p:

$$\begin{aligned} e(p,U_t) &= Min\ [\ p^T x :\ U(x) \geq U_t\ ] \\ &= p^T x_t(p) \end{aligned} \tag{26}$$

Substitute (22) for x_t in (28) to get

$$e(p,U_t) = px^1(p) + p(x^2(p) - x^1(p))\ t \tag{27}$$

noting that $t = U_t$, we can rewrite the above as

$$e(p,U_t) = a(p)\ \theta(x_t) + b(p) \tag{28}$$

where $a(p) = p(x^2(p) - x'(p))$ and $b(p) = px'(p)$. Hence for the expansion set between x' and x^2, the corresponding expenditure function takes the form:

$$e(p,U) = a(p)\ \theta(x) + b(p) \tag{29}$$

The utility function representing a preference ordering R which can be normalized to have a linear expenditure function as in (10) is called quasihomothetic. Such a utility function has expansion paths that are straight lines, but unlike those associated with homothetic preferences, need not pass through the origin. The marginal utility of income associated with such a utility function is independent of income, and is a function only on the n prices. Conversely, a linear expansion path is associated with a linear cost function, which therefore implies that the associated marginal utility of income is independent of income.

2.3 Solving For the Integrating Factor: Linear Expansion Paths: Let U be a utility function representing preferences that are homothetic. Then a normalization of this utility function can be found, such that the corresponding marginal utility of income can be written as a function only of income, $\lambda = 1/m$ (Samuelson 1942) Substituting this in (3), we get

$$EV = m^0 - m^1 \exp \frac{1}{m} \int_{p^0}^{p^1} x(p, m^1)\ dp \tag{30}$$

Consider the utility function representing a preference ordering that is quasihomothetic, with properties described above. Since the marginal utility of income is independent of income; $\partial\lambda/\partial m = 0$, we have the following differential equations as functions of n variables (Seade 1978):

$$\frac{\partial \lambda}{\partial p_i} = - \lambda \frac{\partial x_i}{\partial m} \tag{31}$$

which can be integrated to yield

$$\lambda(p) = \lambda(p^1) \exp\left(- \int_{p^0}^{p^1} \sum \frac{\partial x_i(p^1, m)}{\partial m}\ dp_i \right) \tag{32}$$

which can be substituted into (3) to yield:

Linear Expansion Paths

$$EV = dm + \int_{p^1}^{p^0} \exp\left(- \int_{p^0}^{p^1} \sum \frac{\partial x_i}{\partial m} \right) \sum x_i(p, m^1) \; dp_i \qquad \textbf{(33)}$$

Linear expansion paths which are associated with either homothetic or quasihomothetic preferences, have implications for the marginal utility of income which makes it possible to solve for Hicksian variations from ordinary demand functions. For nonlinear expansion paths, the marginal utility of income is a function of n+1 variables, which makes implicit solutions more complex. McKenzie and Pierce (1976) have shown that the equivalent variation can be obtained in such cases, using a more detailed calculation. Quasihomotheticity, on the other hand, allows a more elegant solution without placing overly restrictive assumptions on the underlying preferences.

II.3

Measuring The Equivalent Variation Exactly: Nonlinear Expansion Paths

---

Integrability theory tells us that the ordinary demand functions give us all the information needed to determine the underlying utility function; hence the expenditure function can, at least theoretically, be calculated from the observed demand functions. However, it was not until recently (Hausman 1981, Vartia 1983), that a practical computational procedure to calculate the Hicksian variations directly from demand functions, without imposing any restrictions on the underlying preferences, has been developed. The basic idea is that the indirect utility function can be recovered from the demand functions if the integrability conditions are satisfied, and the expenditure function can then be obtained by inverting it. Alternately the expenditure function can be calculated directly from ordinary demand functions by considering price and income changes that keeps the consumer on an indifference surface. Since utility remains constant for such budget changes, the compensated income corresponding to these price changes gives the expenditure function corresponding to this indifference surface.

3.1 Calculating The Expenditure Function From Ordinary Demand Functions

Let $\alpha(t) = \alpha(p(t),m(t))$ be a differentiable path in the budget space Ω joining the initial point (p°,m) and the final point (p',m). Let I = [a,b] be the interval describing the change in prices and income. For simplicity, let a = 0, and b = 1; hence for $0 \le t \le 1$, p(t) is the price path joining p° and p', and m(t) is the income path beginning at m°. Let $V(\alpha(t))$ = $V(\alpha(p(t), m(t)))$ be the indirect utility function, whose value changes as we move along $\alpha(t)$, and e(p(t),U(x')) = m(t) be the expenditure or compensation function, defined for the final utility level, which gives the change in income required to maintain the utility level obtained when the individual consumes x' at the final situation. We can differentiate the indirect utility function with respect to t, and using Roy's identity

$$\frac{dV(t)}{dt} = \lambda(p(t),m(t)) \left[\frac{dm(t)}{dt} - \sum x_i(p(t),m(t)) \frac{dp}{dt}\right] \quad \textbf{(1)}$$

It gives the rate of change in utility at every point t, as p(t) and m(t) move from their initial to final values.

Consider the Hicksian demand function h(p(t),U'), which for changes in t keeps the consumer on the indifference surface corresponding to U'. It implies that dV = 0 for all $t \in [0,1]$. Since $\lambda > 0$, the above equation reduces to the following differential equation:

$$\frac{dm(t)}{dt} = \sum x_i(p(t),m(t)) \frac{dp(t)}{dt} \quad \textbf{(2)}$$

where the demand function x(p(t),m(t)) and its derivatives are known, and m(t) is the unknown function which when solved gives the compensating income, e(p(t),U(x')) = m(t) required to keep the consumer on the indifference surface corresponding to U'.

3.2 Solutions To First Order Differential Equations

The demand function $x\colon \Omega \mapsto \Omega^n$ is continuous, and defined for the real $p \in [a,b]$ and all real m. The equation in (2) written as

$$\frac{dm(t)}{dt} = f(t,m(t)) \equiv x(p(t),m(t)) \frac{dp}{dt} \tag{3}$$

is an ordinary differential equation of the first order. A solution of this differential equation is the function $\mu(p(t);p(t),m(t)) = m(t)$, defined as the compensating income. The function μ is of class C^2, and satisfies the following two conditions:

$$\begin{aligned} &(i) \quad (t,\ \mu(t) \in \Omega; \qquad t \in I \\ &(ii) \quad \frac{d\mu(t)}{dt} = f(t,\ \mu(t)); \qquad t \in I \end{aligned} \tag{4}$$

The general solution of the differential equation will involve one arbitrary constant which can be determined if we impose the condition that the solution must satisfy an initial value constraint. The determination of the solution of

$$\frac{dm}{dt} = f(t,m(t)); \quad m(t_0) = m^0, \quad t \in I \tag{5}$$

is the initial value problem at hand. If there exists an interval I containing t_0, and an m satisfying (5), it can be referred to as a solution of (3) passing through (t_0, m°).

EXISTENCE AND UNIQUENESS OF SOLUTION

The question about the existence and the uniqueness of the solution of (3) are answered by an existence theorem referred to as the Picard - Lindelof theorem (see Hartman, 1964).

THEOREM: Consider the function f(t, m) defined on a domain Ω in the n+1 dimensional space. If:

(i) f(t,m) is defined and continuous in Ω for $p^0 \leq p \leq p^1$; $m > 0$

(ii) f(t,m) is locally Lipschitzian with respect to m in Ω, that is, if for a bounded set $\wp$ in Ω, $\exists\ L = L_\wp$, such that

$$|f(t,m^1) - f(t,m^2)| \leq L|m^1 - m^2| \text{ forall } (t,m^1),\ (t,m^2) \in \wp \quad \textbf{(6)}$$

then there exists a unique solution $f(t,t_0,m^\circ)$, and $f(t_0,m^\circ) = m^\circ$ of (3) passing through (t_0,m°).

Further, if the Slutsky matrix associated with the function f(t,m) is symmetric and negative semidefinite, then the solution m(t) is the compensation function $\mu(p; p^*, m^*)$ and $\mu(p^\circ; p^\circ,m^\circ) = m^\circ$.

The existence theorem implies that if the function f(t,m) is continuous in the domain Ω, and with continuous first partial derivatives with respect to m, then there exists one and only one solution passing through a given point (t_0, m°) in $\wp$.

For any $(t_0, m^\circ) \in \wp$ in Ω, let $(y^1(t_0, m^\circ), y^2(t_0, m^\circ))$ be the maximal interval of existence of $f(t, t_0, m^\circ)$ and let Γ be defined by

$$\Gamma = \{(t, t^0, m^0) : y^1(t_0, m^0) < y^2(t_0, m^0);\ (t_0, m^0) \in \rho\} \qquad \textbf{(7)}$$

The trajectory through (t_0, $m°$) is the set of points given by (t,$f(t,t_0,m°)$) for t varying over all possible values for which (t, t_0, $m°$) belong to Γ.

3.3 Closed Form Solutions To Differential Equations

For some differential equations, the solution can be expressed in terms of elementary functions, or in terms of indefinite integrals of such functions. Such a solution is called an explicit solution. Hausman (1981) and Mohring (1971) consider explicit solutions for the case of single price changes. Although generalizations for multiple price changes are not possible, their approach is a useful starting point in understanding more complex solutions of differential equations.

Consider the case of only two goods, where x_1 is the good whose price p_1 changes from some initial value $p_1°$ to p_1' and the price of the other good is normalized to equal one. Let us write the differential equation in (3) as:

$$\frac{dm(p_1)}{dp_1} = x_1(p,m) \qquad \textbf{(8)}$$

From henceforth we shall drop the subscript and it is understood that it is the price of the first good that is being changed; all other variables remaining unchanged. Some of the simplest types of equations can be solved by the computation of integrals. For example, Hausman considers a linear demand function of the form

$$x = ap + bM + dz \tag{9}$$

Thi function can be integrated to give

$$m(p) = c \exp(bp) - \frac{1}{b}\left(ap + \frac{a}{b} + dz\right) \tag{10}$$

where c is the constant of integration. Since U is an ordinal function, we can let c = U', where U' is the final utility level. The above equation can then be solved for the indirect utility function

$$c = V(p,m) = \exp -(bp)\left(m + \frac{1}{b}\,(ap + \frac{a}{b} + dz)\right) \tag{11}$$

which can be substituted into (10) to yield the expenditure function.

Consider another example, a CES function

$$x = \exp(dz)\, p^{a}\, m^{b} \tag{12}$$

This can be integrated using the technique of separation of variables. To do this, the equation can be divided by m^b,

$$\frac{dm}{dp} = \exp(dz)\, p^{a}\, m^{b} \tag{13}$$

and integrated

$$\int \frac{1}{M^{b}}\, dM = \int \exp(dz)\, p^{a}\, dp + c \tag{14}$$

from which the expenditure function can be calculated, using the value of c = V(p',m).

SOLVING DIFFERENTIAL EQUATIONS USING *MACSYMA ODE 2*

More complex specifications of the demand functions, which require involved computations, can be more conveniently handled using a suitable computer software. For example, consider the demand function which is quadratic in prices:

$$x = ap + bp^2 + gm + dz \tag{15}$$

To integrate this, we used MACSYMA, a symbolic differential equation solver. This is the most advanced program available that can be used for explicitly solving ordinary differential equations of the first and second order. It consists primarily of a set of routines based on techniques from Moses' SOLDIER ODE program (Moses, 1967) and Boyce and DiPrima (1969). A copy of the program by Golden, J.P., taken from the proceedings of the 1977 MACSYMA users conference, U.S. Department of Commerce, is enclosed.

The differential equation dm(p)/dp = ap + bp² + gm + dz was solved using MACSYMA ODE 2 to obtain:

$$m(p) = \frac{-b(g^2p^2+2gp+2)}{g^3} - \frac{a(gp+1)}{g^2} - \frac{dz}{g} + V\exp(gp) \tag{16}$$

where

$$V = \exp-(gp)\left(\frac{g^3m + bg^2p^2 + (ag^2+2bg)p + dzg^2 + ag + 2b}{g^3}\right) \tag{17}$$

MACSYMA's ordinary differential equation solver ODE 2 may be used for solving elementary ordinary differential equations: linear, exact, homogeneous, Bernoulli's equation. However, for more complex specifications of demand functions, resulting in a nonlinear differential equations, an analytical solution is often not possible. This is also the case when more than one price changes. We discuss this in some length at the end of this chapter.

3.4 Numerical Solution Of Differential Equations

If an explicit solution of the differential equation is not feasible, it may be possible to solve it using numerical methods. Consider the initial value problem for the first order differential equation in (3) with the initial value $m°$. We seek a solution in the range $0 \le t \le 1$. We assume that the function $f(t, m(t))$ satisfies the conditions in (3), which guarantee that the problem has a continuously differentiable solution, which we shall indicate by $m(t)$.

The idea behind numerical integration is to approximate the solution to (3) at a sequence of points called mesh points. We divide the price interval $[a,b]$ into equidistant intervals through the mesh points $a = t_0, t_1, t_2, ..t_{k-1}, t_k = b$, with the step size $h = (b-a)/k$, such that $t_i = t_0+ih$. The approach of a numerical solution is therefore that of discretization; that is we seek an approximate solution, not on the continous interval $a \le t \le b$, but on the discrete point set $\{t_j;\ j = 0,1..(b-a)/h\}$.

Let m_{tj} be an approximation to the theoretical solution at t_j, that is, to the actual compensated income $m(t_j)$, and let $f_j = f(t_j, m_j)$.

The idea of this algorithm is to move by small steps along the indifference surface corresponding to the final utility, as prices are changed from their initial to their final values. The values of the m_{tj} gives the approximate value of the true compensated income $m(t_j)$ corresponding to p_{tj} and x'.

The computational method that we use for determining the sequence {Mt} takes a linear relationship between m_{tj}, f_{tj}, j = 0....k called a linear multistep method of step number k. A general form of a k - step method may be written as

$$\sum_{j=0}^{k} \delta_j \, m_{tj} \;=\; h \sum_{j=0}^{k} \nu_j \, f_{tj} \qquad \textbf{(18)}$$

where δ_j and ν_j are constants which are arbitrary up to a constant multiplier, since (18) can be multiplied on both sides by the same constant without altering the relationship We can remove this arbitrariness by assuming throughout that $\delta k = 1$. Thus the problem of determining the compensated income m(t) in (3) is replaced by that of finding the sequence {mt} which satisfies the difference equation in (18).

This is done by expressing the solution at t_j, that is $m(t_j)$, in terms of numerical approximations of j-1 successive steps. At each stage of the computation the method will call for the solution of the equation:

$$m_{t_i} = h\, v_i\, f(t_i, M_{t_i}) + \Delta \tag{19}$$

where Δ is a known function of the previously calculated values m_{tj}, f_{tj}, $j = 0, 1....i-1$. It should be pointed out that if the original differential equation in (3) is linear, for example if the demand function is linear, or loglinear, then (19) is also linear in m_{ti}, and there is no problem in solving it. But for more complex forms of demand functions like quadratic, indirect, addilog, which result in nonlinear differential equations, the condition in (6) guarantees that there exists a unique solution for m_{ti}, which can be approached closely by an iteration process. The solution will converge to the actual solution if

$$h < \frac{1}{L|v_j|} \tag{20}$$

where it may be recalled that L is the Lipschitz constant of X with respect to m.

We now turn to the problem of determining the coefficients δ_j, v_j in (18). Any multistep method may be derived in a number of different ways, each of which is primarily an exercise in judicious extrapolation. The principle criterion for choosing any one method over another are ones of efficiency, stability, and accuracy. The whole method of calculation is governed decisively by the number of steps k, wnich also determines the degree of accuracy.

Calculating EV Directly

To explain the method that we used, it would be useful to begin with the simplest numerical method, known as the Euler method. Consider the Taylor expansion for $m(t_j+h)$ about t_j:

$$m(t_j + h) = m(t_j) + h m'(t_j) + \frac{h^2}{2!} m^2(t_j) + \ldots \qquad \textbf{(21)}$$

If we truncate the expansion after two terms and substitute for $m(t_j)$ from the differential equation (3), we have

$$m(t_j + h) \approx m(t_j) + h f(t_j, m(t_j)) \qquad \textbf{(22)}$$

a relation which is in error by

$$0(h^2) = \frac{h^2}{2!} m^2(t_j) + \frac{h^3}{3!} m^3(t_j) + \ldots \qquad \textbf{(23)}$$

Equation (22) expresses an approximate relation to the exact values of the solution of (3). We can also interpret it as an exact relation between approximate values of the solution of (3) if we replace $m(t_j)$, $m(t_j+h)$ by mt_j, mt_{j+h} respectively:

$$m_{t_{j+1}} = m_{t_j} + h f_{t_j} \qquad \textbf{(24)}$$

which is an explicit linear one step method. This is Euler's rule, which is the simplest of all linear multistep methods. The basic idea is that the points (t, m) 'near' the solution curve of the problem are calculated successively by (22). Hence the solution is obtained by integrating the differential equation over the interval (t,t+h) and then the integral of f(t,m(t)) dt is replaced by an approximation

$$m(t+h) - m(t) = \int_{t}^{t+h} f(t, m(t))\, dt \tag{25}$$

The error associated with it in (23) is called the local truncation error, which for this method is $O(h^2)$, and is equal to zero if the solution of (3) is a polynomial of degree not greater than one.

The above method was a single step integration, where the numerical solution was obtained solely from the differential equation and the initial value. This is a naive method of limited accuracy, because it requires the value at one mesh point to compute the value at the next. Once the values at a number of points of the numerical approximation have been computed, they could be used to aid in the computation of later points. A multistep method does exactly this; and hence is a more efficient method providing a more accurate solution. We therefore chose a multistep method, called the *predictor - corrector* method to solve our differential equation in (3).

The algorithm consists of two parts: first, a starting procedure in which an approximation to m_{t1} is computed from m_{t0} by linear extrapolation, providing $m_0 \ldots\ldots m_{i-1}$, which are the approximation to the exact compensated income at the points $t_0 \ldots t_{i-1}$. This is called the *prediction*. Since this process does not make use of the differential equation in any way, a *correction* process is applied to the approximate values.

Calculating EV Directly

The method becomes efficient and keeps the number of iterations to the minimum, if the initial starting values can be made as accurate as possible. We did this using an explicit difference method, called the Adams - Bashforth method, to estimate fairly reasonable values for m_{tj}. This value is used as the initial estimate, which is then corrected using an implicit formula by iteration, until there was convergence.

This process is applied recursively based on the numerical approximations of the k successive steps, to compute $m(t_k)$,.... We repeated this process for 1000 iterations until there was no further change in the predicted value of m, which indicated convergence.

3.5 A Numerical Example

We illustrate the above algorithm using the demand function estimated by Parks R.D. and D.Weitzal from the data generated by the Wisconsin Time-of-Day electricity pricing experiment. The experiment was designed to distinguish four time periods in a typical day with x_i, i = 1..4, denoting the quantity demanded in each time period, and p_i the associated price of electricity for that period.

Homothetic separability was used as a maintained hypothesis in this study, making the quantity demanded in any time period a function only of electricity prices and the total electricity expenditure. The authors estimated a generalized Leontief model of demand, and was of the following form:

$$x_i = \frac{m \sum_j d_{ij}\left(\frac{p_j}{p_i}\right)^{\frac{1}{2}}}{\sum_i \sum_j d_{ij}\, p_i^{\frac{1}{2}}\, p_j^{\frac{1}{2}}} \tag{26}$$

with $d_{ij} = d_{ji}$; $\Sigma_i \Sigma_j d_{ij} = 1$. The movement from a flat rate to a time differentiated tariff can be represented by the change in the price vector $p° = (p_1°, p_2°, p_3°, p_4°)$, with $p_i° = 4$ cents/Kwh to $p' = (p_1', p_2', p_3', p_4')$. For purposes of illustration we use 8, 4, 2, and 1 cent as the prices for the peak, shoulders, and offpeak periods respectively.

The change in the consumer's welfare, as measured by the equivalent variation, was calculated using the differential equation in (3). The initial condition used for this problem was the average monthly expenditure on electricity, which was estimated by the authors to be \$ 26.82.

A program called ISLM ROUTINE - DGEAR was used to find approximations to the solution of the first order differential equation in (3), with initial conditions $m(t_0) = 26.82$. Two classes of implicit linear multistep methods were used in the process: the first is the implicit Adams methods (described above) up to order 12, and the second is the backward differentiation formula, up to order 5 (See Gear, C.W., 1971 for details). The demand function in (26) with values of the coefficients (d_{ij}) estimated by the authors was used in the subroutine. A multistep predictor - corrector method was used, whose order was automatically chosen by the subroutine as the integration proceeded. The starting method is automatic, and the information retained by the program about the previous steps was stored in such a way as to make the interpolation to a nonmesh point straghtforward.

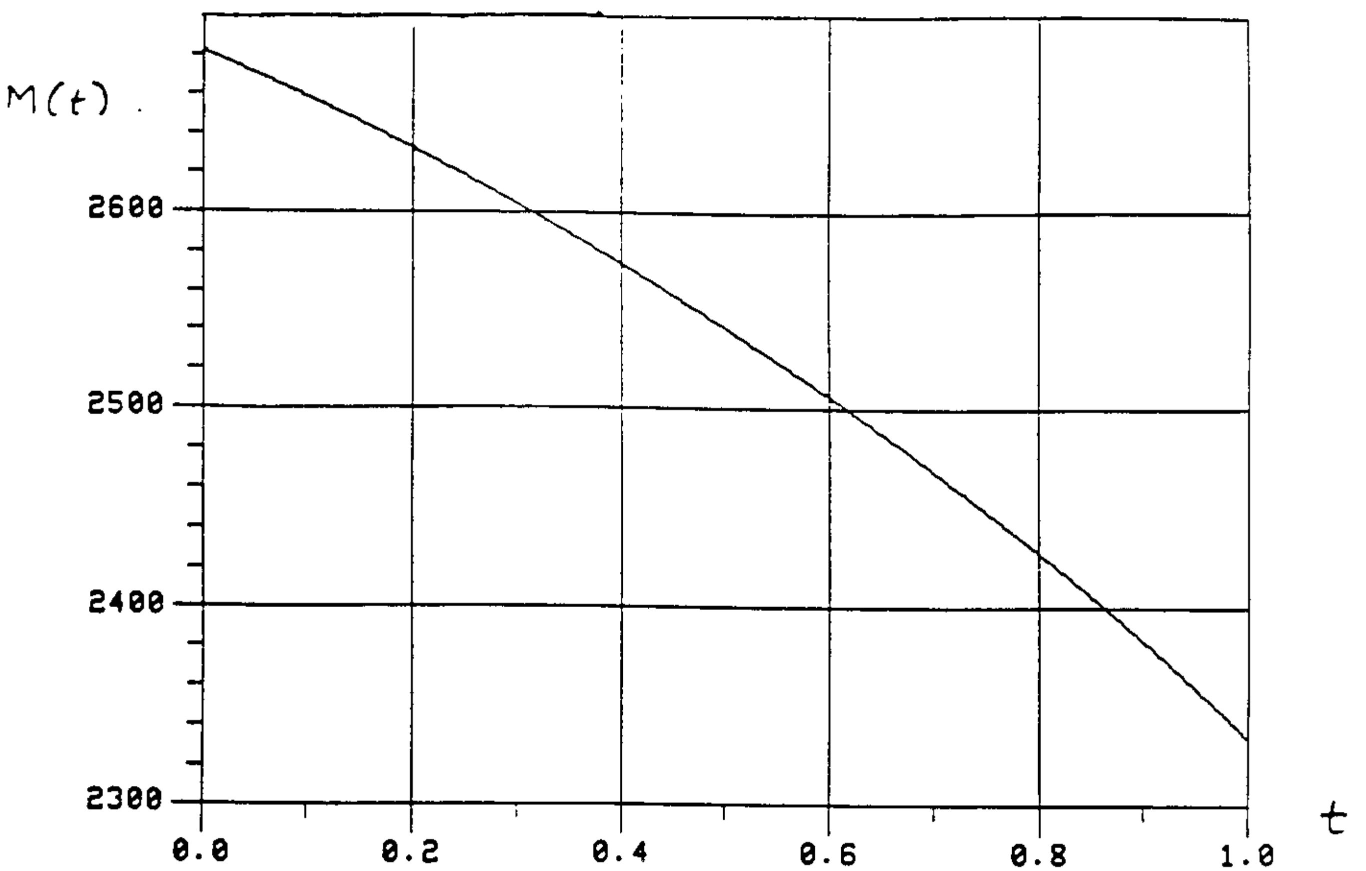
We divided the interval [0,1] into 1000 steps. The distance between two mesh points had a step length h = 0.0001, and the error tolerance was set at (0.000001). Since the Lipschitz condition was satisfied, the step length used was adequate for the convergence of m_{tj} to $m(t_j)$. The value of m_t was found to be 23.335, showing a surplus gain of $3.485. The values of the compensated income estimated at each mesh point was plotted, and is enclosed.

3.6 Calculating The Equivalent variation and Marshallian Surplus: A Comparison Of Computational Ease

Despite the well known theoretical problems underlying the use of the Marshallian surplus, actual practitioners of cost-benefit analysis continue to use it in the belief that it is easier to calculate the Marshallian surplus than the equivalent variation, requiring merely the integration of ordinary demand functions with respect to the price changes. As shown in the above sections, the availability of advanced software has made the computation of the equivalent variation from ordinary demand

$$\frac{dM(t)}{dt} = \sum_i x_i(P(t), M(t)) \frac{dP(t)}{dt}$$

$$x_i = \frac{M \sum_j d_{ij} \left(\frac{p_j}{p_i}\right)^{\frac{1}{2}}}{\sum_i \sum_j d_{ij}\, p_i^{\frac{1}{2}}\, p_j^{\frac{1}{2}}}$$



functions feasible.

Calculating the Marshallian surplus may not always be easy, often requiring the use of advanced computing software. While it is true that for certain kinds of demand functions like linear, or loglinear, the integral can be expressed in terms of elementary functions, it may not be possible to do so for more complex specifications of demand functions. Examples of this are the frequently used demand functions associated with the addilog, translog, and other utility functions that are not separable in income. Even demand functions obtained from additive utility function that are linear in income like the members of the Bergson family pose considerable difficulties for integration.

The problem in computing the Marshallian surplus arises when the demand function takes the form $x = g(p_i)/h(p_i)$, and we require to integrate it with respect to the changes in p_i. Let the demand function $x(p_i)$ be a rational function with a domain in (p,m); the objective is to express the integral of x in terms of (p,m, log).

Without loss of generality, let $g(p_i)$, $h(p_i)$ be coprimes, and $g(p_i)/h(p_i)$ be a proper fraction. Basically the method is to split the integrand into the sum of simple functions, which can be integrated directly. That is, we have to break up the function x into parts which are easy to integrate, with the restriction that if x is integrable, then each of the parts that we split it into must likewise be integrable. An example where this could go wrong is given by :

$$\frac{1}{2}\exp(p_i+1)^2 = \int (p_i+1)\exp(p_i+1)^2 = \int p_i \exp(p_i+1)^2 + \int \exp(p_i+1)^2 \quad \textbf{(27)}$$

However $\exp(p_i+1)^2$ has no elementary integral.

The partition that is applied to x is the natural one of splitting it into a polynomial and proper fraction. Then from Liouville's theorem, the function x has an integral of the following form:

$$\int x = y_0 + \sum_j c_j \log y_j(p_i) \quad \textbf{(28)}$$

where y_0 is a rational function in p_i, and $y_1(p_i)\ldots.y_k(p_i)$ are square free coprime polynominals, and the c_j's are constants. The problem in integrating the function x arises due to the presence of p_i in the denominator of x, and hence also in the denominators of y_0 and the y_js. Let us consider each of these terms. We can deduce some facts about the denominator of y_0 by differentiating the equation above. First, notice that the derivative of $c_j \log y_j(p_i)$ is $c_j\, y'_j(p_i) / y_j(p_i)$. Since $y_j(p_i)$ is square free and thus coprime to its derivative, it therefore has a square free denominator. Hence the derivative of the sum of the logarithms is a rational function with square free denominator; in fact, the denominator will be precisely the product of the arguments of the logs. Now consider the term y_0. Differentiating this shows that any factor appearing in the denominator of y_0 will appear in the denominator of x to a higher power and we know that the power is exactly one greater. In other words, given x we can compute the denominator of y_0 directly from the denominator of x.

This leads us directly to the basis of the method behind integrating the function x. Given a function $x_i = g(p_i)/h(p_i)$, we need to factored $h(p_i)$ into linears, such that

$$h(p_i) = (p_i - c_k)^k \tag{29}$$

which will then allow us to write $g(p_i)$ / $h(p_i)$ into partial fractions:

$$\frac{g(p_i)}{h(p_i)} = \sum_{j=1}^{k} \frac{g_j}{(p_i - c_k)^{e_j}} \tag{30}$$

Then each term can be integrated independently:

$$\int \frac{g_{j1}}{(p_i - c_j)} = g_{j1} \log (p_j - c_j) \tag{31}$$

Of course not all functional forms of demand need to have the denominator factored. For example, consider the demand function corresponding to the Linear Expenditure System (LES).

$$x_i = c_i - \frac{\beta_i}{p_i} \sum_j p_j c_j + \frac{\beta_i}{p_i} m \tag{32}$$

To integrate this function with respect to p_i does not require us to factor the denominator, since we can integrate in parts. For example if there are three goods i=1..3, the demand function for x_1 can be integrated with respect to p_1 to get

$$-[\log p_1(c_2p_2 + c_3p_3) + c_1p_1]\beta_1 + m\log(p_1)\beta_1 + c_1p_1 \qquad \textbf{(33)}$$

In general, the situation will not be as convenient. For example, consider the demand function obtained from the indirect addilog function:

$$x_i = \frac{\alpha_i\beta_i\left(\frac{m}{p_i}\right)^{1+\beta_i}}{\sum_j \alpha_j\beta_j\left(\frac{m}{p_j}\right)^{\beta_j}} \qquad \textbf{(34)}$$

The denominator contains the term p_i to the power of β_i. We know that the only restriction that we can place on the β_is is that $\beta_i \geq -1$ for all i, with the equality holding for at most one commdity (see Parks 1969, and Yoshihara 1969). Since the β_i are in general nonintegers, we have a complex variable in the denominator. Complex variables have nonunique roots; hence it is impossible to factorize the term linearly. We cannot express the integrals of such functions as closed forms, and can only be integrated numerically. Calculating the Marshallian surplus in such cases is not any easier than calculating the equivalent variation. Given that the latter can be computed from knowledge only of the ordinary demand function, the change in consumer surplus can be measured exactly.

An Application: Measuring Consumer Welfare Effects Of Time Differentiated Electricity Rates

---

The literature on surplus measures has traditionally focused on situations where changes in the consumer's budget lead to monotonic changes in utility. Our objective here is to consider those paths in the budget space along which utility varies non monotonically, and to show that the standard results of surplus measures derived for monotonic utility changes may not apply in these situations. Budget changes that lead to non monotonic changes in utility are not uncommon: price increases are often accompanied by lump sum transfers; a subset of prices may be changed such that some prices are increased, and others decreased. An example of the latter kind, which has received considerable attention recently, is the proposed change of electricity prices from the current flat rates to peak load or time differentiated rates. Typically, such a change would involve an increase in the price of electricity for the peak periods with a simultaneous reduction in off-peak prices. If rectangular paths are used to describe such changes, we have a situation where the price changes lead to nonmonotonic changes in utility. Hence for purposes of illustration, we consider the problem of measuring consumer surplus, when electricity prices are changed from a flat to time differentiated rates.

Several projects have been conducted to study the effects of time-differentiated rates in the U.S. (see D.J.Aigner 1984).

Application: Electricity Pricing

Many of these studies (for example, Parks and Weizel 1984), use the Marshallian surplus to measure the changes in consumer welfare. It is argued that since a consumer's expenditure on electricity is only a small part of his income, the Marshallian surplus can be used to approximate the true surplus change. We will show that this assumption does not hold, since these are paths describing nonmonotonic utility changes. The Marshallian surplus does not approximate the equivalent variation closely, even when the income elasticities and expenditure proportions are small.

Some effort to measure the changes in consumer welfare in terms of the Hicksian variations has been done (Caves, Schoech, Christensen & Hendricks, 1984; Howrey and Varian, 1984; Gallant and Koenker, 1984). They use the more common approach of first specifying the utility function, and then estimate the unknown parameters from the derived demand functions. Our approach differs from their work in that we begin with the observed demand function and then derive the expenditure function from it. This is considered to be the preferred approach, since it does not force the demand data to fit a pre-specified utility function (Hausman, 1981). Further, this method can be applied for any specification of the demand function as long as the integrability conditions are satisfied. No restrictions need be placed either on the preferences of the consumer, or on the paths of integration.

Application: Electricity Pricing

We begin with a brief background on the rate structures and energy policy in the United States and the present crisis facing the energy sector. We then look at the different demand functions estimated under the time-of-day experiments, and the assumptions underlying their specifications. Using the data from the peak-load experiment conducted in Connecticut, we show that the changes in the consumer surplus due to the proposed change can be measured exactly using the equivalent variation. We do this using only the observed demand functions estimated by Lawrence and Braithwaite (1979). Given that the underlying preferences corresponding to these demand functions were non-homothetic, we also found that the Marshallian surplus diverged substantially along alternate paths. Further, our calculations showed that the Marshallian surplus failed to approximate the equivalent variation closely.

1. The Problem Of Electricity Pricing

Until the early seventies, the electric utilities lived in so favourable a world, that the problem of designing optimal rate structures in terms of marginal cost pricing could be ignored in the United States and Canada. Several unusual conditions existed at that time that led to decreasing costs: technical innovations in generation and transmission of electric power, extensive scale economies as generating facilities grew. Fuel costs were stable and low with little concern about future supplies. Apart from this, generating power sites were not scarce, since the aggregate emissions of thermal and air pollutants were not considered significant social costs. Most utilities had excess capacity; reserve margins were about twenty five percent above peak loads. Marginal costs even in the peak periods were therefore low, and the early rates, such as declining block rates, were aimed at promoting the consumption of electricity to allow utilities to benefit from the economies of scale.

It took an abrupt reversal of these conditions to create a critical public attitude towards traditional utility pricing. Since 1970, scale economies reached their limit in generation. Fuel costs began to rise rapidly with the energy crisis of the seventies. Not only did the concern of nuclear hazards and pollution become major hindrances to electricity growth, but have decreased the potential for more cost reductions through further technical innovations. The attitude now is to conserve capacity rather than to enlarge it, and marginal cost pricing is receiving increasing attention.

Although electric power is physically a homogeneous commodity, its value to consumers is not the same at different times of the day or during different seasons of the year. Hence the demand for electricity varies at different times in terms of the level of industrial and human activities, and the need for weather control. A typical daily load curve shows that the demand for electricity peaks for a few hours, and for any utility, the timing of peak demands follows regular patterns.

The cost of supplying electricity can be divided broadly into operating costs, and the cost of resources used to provide capacity. Since electricity once generated cannot easily be stored, to supply electricity with some level of reliability, utilities have to maintain an amount of capacity depending on the demand in the peak periods. Hence, the peak periods are associated with larger invested capital costs than off-peak.

Although capital costs often constitute the major portion of the total costs, the cost of fuel is also quite significant. The opportunities for trading off capital costs for fuel costs depends in great part on the technology used by the utility in terms of the primary energy used to generate the power. The bulk of the power in the United States is generated by thermal units which use fossil fuels or nuclear energy to drive the turbines. Nuclear plants have high capital costs but low fuel costs, compared to gas turbines, which have relatively high running costs but low capital costs. Generally, utilities have a combination of generating plants using different fuels. As demand increases during peak periods, the use of fuel intensive units increase. Hence peaks in the load curve are associated with both increased operating and capital costs.

Given the differences in the marginal costs between peak and off peak periods, significant savings are possible, if the demand in the peak periods can be reduced or shifted to off-peak periods. One way to achieve this is by establishing time-of-use rate structures which will reflect the cost of supplying electricity in different periods. If the price in each period is set at the marginal cost of its supply, it will ensure that each customer will use only that amount of electricity whose value to him is at least equal to the costs of supply. Such a price structure is expected to increase energy conservation by encouraging consumers to choose more efficient use of energy in terms of the load factor, the choice between alternate sources of energy, as well as their choice of capital equipment.

In spite of this, there is considerable caution on the part of utilities and regulators towards tariff revisions. Although time differentiated pricing promises benefits in the form of elimination of certain welfare losses resulting from current failure to price optimally, there will be substantial metering costs in administrating a more detailed tariff under peak-load pricing. A move to time-of-day (TOD) rates requires investment in time-of-use meters, the costs of which varies between $80 and $160, depending on particular features. Acton and Mitchel (1980) assume an installed cost per meter of $150. With an estimated life of 15 years, and an 8 percent discount rate, they estimate a monthly cost of $1.42.

Given this cost of implementing time-of-day rates, a move to such a system of pricing will be worthwhile only if the net benefits to society exceeds this cost. For a utility's largest customers such as the industrial sector, these costs of implementing peak load tariffs are only a small fraction of their total costs of electricity, and hence they gain from peak-load pricing. However in the case of of smaller customers like residential users, the size of these gains relative to the costs of implementation is uncertain. Hence whether time differentiated prices should be implementated or not depends critically on measuring the benefits of the change, and then comparing these gains to the costs of implementing them. To quote D.J.Aigner, "the issue to be addressed at the present moment is in a word, welfare. For residential customers - do the benefits to society of placing households on TOD rates out weigh the costs of implementation" (Journal of Econometrics 1984).

2. *Residential Demand For Electricity*

The consumer's demand for energy, whether electricity, gas, or oil is derived from the demand for the services it can perform. The object of choice is the services of a stock of appliances that use electricity as an input. Since appliances are durable goods, there is accordingly a well defined distinction between the short run, during which the stock of appliances is fixed, and the long run during which it can be varied. Here, we are concerned only with short run demand for electricity, which can be viewed as the choice of a utilization rate of the existing fixed stock of appliances.

Application: Electricity Pricing

In considering the short run demand for electricity under TOD rates, the day is divided into a series of periods $t_1 \ldots t_n$. The electricity consumed in each of these periods is treated as a distinct commodity, denoted by the vector $x = (x_1 \ldots x_n)$. Let $p = (p_1 \ldots p_n)$ be the price of electricity in each period. It is assumed that, apart from electricity, there is also a vector of non electricity goods $z = (z_1 \ldots z_m)$, whose prices may be denoted by $q = (q_1 \ldots q_m)$.

Consider a household whose preferences can be represented by a utility function:

$$U = U(x,\ z) \tag{1}$$

The household is assumed to choose elements of x and z, in order to maximize the utility function in (1), subject to a total expenditure or income constraint:

$$m = \sum_{i=1}^{n} p_i x_i + \sum_{j=1}^{m} q_j z_j \tag{2}$$

SEPARABILITY OF THE DEMAND FOR ELECTRICITY

The TOD experiments conducted by the various groups were designed to provide detailed information about the response of residential demand to time differentiated rates. Since the data from such experiments does not permit complete modeling of the household's choices with respect to all the commodities, these models assume that the demand for electricity is separable from all other goods, and implies the following weakly separable direct utility function:

$$U = U(F(x), z) \tag{3}$$

where U(.) and F(.) are continuous, and U(.,Z) is strictly increasing. The function F(.) is viewed by Parks and Weitzel (1984) as a household production function that combines a predetermined stock of capital and electricity as inputs to produce F. Consumers therefore make decisions about the cost minimizing input combinations of capital and electricity for producing F. These determine the unit cost of the produced commodity, and consumers then determine the optimal level of consumption of X, as well as the other goods and services.

Separability of the utility function allows a two level budgeting procedure. This approach implies that the household decides on the total expenditure on electricity along with the amount to be spent on other goods. This decision uses a price index for electricity consumption at different times of the day along with the prices of all other goods. Given the expenditure on electricity, the share of expenditure allocated to each time period is then decided upon, depending on the relative prices of electricity. Partitioning the analysis into two stages allows us to focus only on the electricity prices and expenditure in the second stage, and disregard the effects of all other variables.

HOMOTHETICITY OF THE SUBUTILITY FUNCTION

Estimation of the electricity demand equations requires the specification of the functional form for F(.). Most studies assume the subutility function to be homothetic. For example, Caves and Christensen (1980) and Howrey and Varian (1984) use a CES specification. Atkinson (1979), and Hausman, Kinnucan, and McFadden (1979) assume a translog, and Parks and Weitzel (1984) estimate demand functions based on the generalized Leontief system.

The homothetic restrictions may be interpreted as requiring ceteris paribus, that the share of expenditure on electricity in each time period be independent of the total expenditure on electricity. It is well known that this places strong a priori restrictions of unitary income elasticities and expansion paths that are not only linear, but pass through the origin. Whether such restrictions are justifiable depends on whether homothetic separability is consistent with observed behaviour.

In a recent article by Kohler and Mitchel (1984), it has been pointed out that the assumption of homothetic separability for demand under TOD rates has not been tested. While these assumptions of homotheticity and separability can in theory, be statistically tested, the TOD experiments that were conducted were such that either the experimental design did not allow it, or the results were ambigous. Imposing homothetic separability as an untested hypothesis carries the risk of misspecification and restrictions that do not reflect true consumer behaviour.

While separability is necessary for consistent multistage budgeting, homotheticity is not (Barnett, 1983). This is because the role played by homotheticity in multistage budgeting can be satisfied by quasihomotheticity, which is a more reasonable assumption to make. The translation from homothetic to quasihomothetic preferences is easily accomplished for any demand model. All that is needed is an affine translation of the origin of the consumption space to a 'committed consumption bundle' (see Deaton and Muellbauer, 1980).

Whether the subutility function is assumed to be homothetic or nonhomothetic has important implications for the measurement of consumer surplus, as discussed earlier. If the function is assumed to be homothetic, then the Marshallian surplus can be validly used to measure the surplus change. However if it is a nonhomothetic function, then only the equivalent variation is a valid surplus measure. The Marshallian surplus is not only path dependent, but may vary considerably from one path to another, since these are nonmonotonic with respect to utility changes. Further, it may may fail to provide a good approximation to the equivalent variation.

3. *Measuring Consumer Surplus Under TOD Rates: A Review*

The question of whether on strict cost benefit grounds, TOD pricing for residential consumers will be a worthwhile undertaking has been raised (Aigner, 1984), but still remains unanswered. Several cost benefit studies have been conducted using different case studies, but differ in their conclusions, mainly due to the differences in the cost-benefit methods used (Caves et al, 1984). Given that the decision to implement TOD rates depends on whether society benefits by such a change, the importance of measuring the cost and benefits accurately can be scarcely overemphasized.

To avoid the problem of path dependence of the Marshallian surplus, Turvey has suggested the use of a non rectangular path, where all prices are changed proportionately (Turvey 1976). The surplus change for any time period is measured as the area left of an *interpolated curve*. To obtain these curves, the quantity consumed under flat and TOD rates ${x^0}_i(P^0)$ and ${x^1}_i(P^1)$ are first calculated. These consumption levels and the associated prices are used to draw a linear demand function, obtained by joining points $({x^0}_i,P^0)$ and $({x^1}_i,p^1)$ for each period i. The slopes of these interpolated curves equals $\partial x_i \;/\; \partial p_i + \partial x_i \;/\; \partial p_j$, i,j = 1....n; i ≠ j. Turvey suggests the approximation of the area left of this interpolated cure by a first order Taylor series expansion, which yields

$$\sum_i \Delta p_i \, x_i + \frac{1}{2} \left(\Delta p_i \Delta x_i\right) \tag{4}$$

Such an approach assumes that the demand functions are locally linear and will differ from the true value to the extent that the demand functions are nonlinear.

While the values of Δp_i are known, there still remains the question of calculating the values of Δx_i. Some of the studies (Wenders and Taylor 1976; Acton and Mitchel 1980) obtain these values from price elasticities. This approach assumes that the demand functions have elasticities that are locally constant. To the extent that they are not, these elasticities can give only an approximate value of the x_i s. The limitations of using a summary statistic like elasticities to calculate the impact of TOD rates on load curves has been pointed out by Malko and Faruqi (1980), "....even a complete set of point elasticity estimates will not generate accurate predictions of load shape modifications, except in cases where the underlying consumer's preferences are characterized by constant elasticity demand functions. In general it will be preferrable to utilize the full system of demand functions to predict the impact of TOD rates on utility loads rather than a set of point elasticity estimates."

The other approach is to calculate these values of x_i from the demand functions. Such a method requires information on the demand functions in each period, as functions of prices in all periods. Then the old values of the prices of electricity and the rates under TOD can be used to calculate the change in the quantities demanded in each period.

There has also been some attempt to measure the Hicksian variations (Caves et al 1984; Parks and Weitzel 1984). The approach of Caves et al is to specify the utility function first and then estimate the unknown parameters from the derived demand functions. Since the only observed data are demand data on econometric grounds, it is considered preferable to fit the demand curve, and not the utility function, to observed data.

In their approach, Parks and Weitzel assume that the subutility function F is homogeneous, and hence the corresponding indirect utility function is written in a separable form as $m\Phi(p_1 \ldots p_n)$, where $m = \Sigma_i \; p_i \; x_i$ is the sum of the total expenditure on electricity, and Φ is interpreted as the inverse of a price index. This allows the cost or expenditure function to be separable into a price and quantity index. The welfare change is then calculated by measuring the effect of a change in the aggregate electricity price on the aggregate quantity. Such a procedure requires that the consumer preferences be homothetic, an assumption that is highly questionable.

The method that we use differs from the above, in that we begin with the ordinary demand functions, and then measure the consumer surplus exactly in terms of the equivalent variation, without any resort to approximations. This is done without placing any restrictions on the preferences, or on the paths of price change. The method is general, and can be applied for any kind of multiple price changes, and for any specification of the demand function, provided they satisfy the integrability conditions.

4. Measuring Exact Consumer Surplus Using Data From Connecticut TOD Experiments

For purposes of illustration, we use the results of the Connecticut Peak Load Pricing Test, one of the projects funded by the U.S. government to study the effects of TOD pricing on demand. The experiment involved the customers of Connecticut Light and Power Company; a Northeast Utilities subsidiary. Using this data, Lawrence and Braithwaite (1979) have estimated demand functions for electricity by time of day which corresponds to the Linear Expenditure System. Our reasons for choosing this functional form is because such a demand system implies quasi homothetic preferences. The corresponding expansion paths are linear, but are not forced through the origin. Linearity of expansion paths are sufficient to allow for the two level budgeting procedure described above.

The data used to estimate the LES demand system are from the Connecticut peak load pricing test from July 1974 through December 1976. Demand was integrated into six time-of-day blocks corresponding to the TOD rate schedule:

1.	off peak	9 P.M.	to	8 A.M.
2.	normal	8 A.M.	to	10 A.M.
		noon	to	1 P.M.
		3 P.M.	to	9 P.M.
3.	peak	10 A.M.	to	noon
		1 P.M.	to	3 P.M.

Application: Electricity Pricing

Meters were installed in approximately 200 households and readings of electricity consumption was recorded every fifteen minutes.

Prior to the TOD experiment, customers faced a declining block rate. However for purposes of illustration, we will assume an average flat price of of 4 cents per Kwh. The TOD rates used for the experiment were 1, 3, and 16 cents for the offpeak, normal, and peak periods respectively.

The authors estimated demand functions corresponding to the Linear Expenditure System, which took the following form

$$x_i = b_i + \frac{a_i}{p_i}\left(m - \sum_k p_k b_k\right) \tag{5}$$

where x_i refers to the quantity of electricity consumed in period i per day.

The demand for electricity in each period also depends on other variables such as the stock of appliances, the weather, demographic and socioeconomic variables (see Acton and Mitchel 1980). The study by Lawrence and Braithwaite incorporate these variables into the committed quantities b_i, making them linear functions of the characterestics of the households. These characterestics include appliance stock, demographic and socio-economic characteristics of the household, and other functions of these variables including iterative terms. The values of the parameters a_i and b_i are then estimated. We estimated the values of the parameters from information of the elasticities provided by the authors:

$$a_1 = 0.0763 \qquad b_1 = 9.79406$$

$$a_2 = 0.2478 \qquad b_2 = 8.77126$$

$$a_3 = 0.6759 \qquad b_3 = 4.12840$$

The total expenditure on electricity under the TOD rates was found to be $1.2043 per day. In the study by Hausman et al, (1979) for Connecticut, they found that the total expenditure under TOD rates declined by 5%. Hence we took the initial expenditure on electricity to be $1.2677.

The effect of the change in prices from a flat rate to the TOD rates on the consumer can be measured using the equivalent variation. We used the method described in chapter II.3, which involved a numerical solution of the differential equation

$$\frac{dm(t)}{dt} = \sum_i x_i(p(t),\, m(t)) \; \frac{dp_i(t)}{dt} \tag{6}$$

where x_i is as given in (5). The differential equation was solved numerically using the Adams - Moulton Predictor Corrector method, using the algorithm DIFSUB 407 by C.W.Gear. Given the initial value $m° = 1.2677$, the value of the equivalent variation was calculated to be - $ 0.2449.

5. *Comparing The Marshallian surplus With The Exact Measure*

In some of the studies dealing with the measurement of the welfare effects of a change to TOD rates, use is made of the

Marshallian surplus, in the belief that approximation errors would be small (Parks and Weitzal, 1984; Caves, Christensen, Schoech and Hendricks, 1984). The justification for this is the argument that the consumer's expenditure on electricity is only a small proportion of the consumer's total income. For example, Parks and Weitzel calculate the size of the budget share for their data to be 0.016, and the income elasticity to be 0.229, and hence conclude that the use of the Marshallian surplus would provide a good approximation to the equivalent variation.

We calculated the Marshallian surplus for the above price changes, for the LES demand function. Using rectangular paths, we calculated the Marshallian surplus along alternate paths, and found that the values varied significantly. The range of values varied from $ - 0.2147 to $ - 0.4852; a substantial difference of 0.2704.

Taking a path along which the prices of the offpeak, normal, and peak period prices are changed consecutively, the value of the Marshallian surplus was found to be $ - 0.4735. Hence the percentage deviation of the Marshallian surplus from the equivalent variation is therefore a significant 48.28% .

We also tried Turvey's method using the interpolated demand curve. The change in quantities was calculated exactly using the demand functions in each time period, and not through price elasticities. The change in consumer's surplus was found to equal 0.4310, showing only a marginal improvement, since the percentage deviation from the equivalent variation is still a substantial 43.18%.

Rectangular paths used to describe multiple price change, where price increases are mixed with price decreases, are ones along which utility varies non monotonically. In such cases, only the equivalent variation can be validly used to measure changes in consumer surplus. The Marshallian surplus is not only path dependent, (if the underlying preferences are nonhomothetic), but diverges substantially along alternate paths. Moreover, it fails to approximate the equivalent variation within small errors.

CONCLUSIONS

The traditional approach to the problem of measuring consumer surplus has been to focus only on those situations which result in monotonic changes in utility. This can be seen from the fact that the literature has mainly dealt with single price changes, or even when multiple price changes have been considered, it has been one where all prices move in the same direction. Since the kind of paths commonly used in practical applications are rectangular, these are clearly situations where utility varies monotonically.

However, there are many practical situations where policy changes result in paths in the budget space along which utility varies nonmonotonically. Measuring changes in consumer welfare in these situations is far more complex, and the problems that arise are quite distinct from those that are encountered for monotonic paths. Hence a detailed analysis of the theoretical and practical issues in measuring consumer surplus for non - monotonic paths, from ordinary demand functions, is warranted.

The integrability conditions in terms of the symmetry and negative semidefiniteness of the Slutsky matrix, provide the necessary conditions for constructing a surplus measure from demand functions. The Marshallian surplus and the Hicksian variations can be expressed as a line integral which measures the areas left of the appropriate demand functions, between the price lines. However, the integrability conditions are not sufficient for the Marshallian surplus to be a valid surplus measure. Some additional restrictions either on the underlying preferences, or on the paths of integration have to be imposed for the Marshallian surplus to be a valid measure.

While this is a well known problem, it can usually be ignored for paths along which utility varies monotonically. This is because for monotonic paths, the Marshallian surplus moves in the same direction as the utility change, and hence can at least provide an ordinal measure of utility change However, this no longer holds for nonmonotonic paths. In the absence of constancy of the marginal utility of income, the Marshallian surplus need not bear the same sign as the utility change, and hence can give an incorrect ranking.

The Marshallian surplus has the added problem of being dependent on the path of integration, unless the marginal utility of income is constant along the path. Again, this does not create any serious problem for monotonic paths, since the Marshallian surplus varies only slightly from path to path between two end points. For nonmonotonic paths, the dependence of the Marshallian surplus on the path of integration can be a serious problem, resulting in substantial divergence in the values of the Marshallian surplus.

It is well known that the compensating variation is not a valid index of utility change, which does not permit it to correctly compare several situations. This creates problems in the use of the compensating variation for budget changes that lead to nonmonotonic changes in utility. Like the Marshallian surplus, the compensating variation can give rise to incorrect rankings of situations for nonmonotonic paths.

Only the equivalent variation is a valid index of surplus change for budget changes that lead to nonmonotonic changes in utility. However, the equivalent variation is measured as the areas left of the Hicksian demand functions, which are not generally observed. While for monotonic paths, the equivalent variation can be reasonably approximated by the Marshallian surplus, it is no longer possible to do so for nonmonotonic paths. In these situations, the Marshallian surplus fails to approximate the Hicksian variations closely, even in those situations where it would have for monotonic paths.

The equivalent variation however, can be calculated from just the ordinary demand functions. No restrictions need be placed either on the underlying preferences. The equivalent variation can be calculated for any specification of the demand function, provided the integrability conditions are satisfied. Hence the consumer surplus can be calculated exactly even for nonmonotonic paths.

APPENDIX I: MACSYMA ODE 2

The MACSYMA code for ODE2 follows. (This code comes from the file JPG;ODER 27. Certain less important sections have been omitted.)

```
/* The Ordinary Differential Equation Solver.
   This package consists primarily of a set of routines taken from Moses'
   thesis and Boyce & DiPrima for solving O.D.E.s of 1st and 2nd order.
   The top-level routines are ODE2, IC1, IC2, and BC2.  */

ODE2(EQ,YOLD,X):=SUBST(YOLD,YNEW,ODE2A(SUBST(YNEW,YOLD,EQ),YNEW,X))$

ODE2A(EQ,Y,X):=BLOCK([DE,A1,A2,A3,A4,Q],
   INTFACTOR: FALSE, METHOD: 'NONE,
   IF FREEOF('DIFF(Y,X,2),EQ)
     THEN IF FTEST(ODE1(EQ,Y,X)) THEN RETURN(Q) ELSE RETURN(FALSE),
   IF DERIVDEGREE(DE: EXPAND(LHS(EQ)-RHS(EQ)),Y,X) # 2
     THEN RETURN(FAILURE(MES1,EQ)),
   A1: COEFF(DE,'DIFF(Y,X,2)),
   A2: COEFF(DE,'DIFF(Y,X)),
   A3: COEFF(DE,Y),
   A4: DE - A1*'DIFF(Y,X,2) - A2*'DIFF(Y,X) - A3*Y,
   IF PR2(A1) AND PR2(A2) AND PR2(A3) AND PR2(A4) AND
      FTEST(HOM2(A1,A2,A3,Y,X))
     THEN IF A4=0 THEN RETURN(Q) ELSE RETURN(VARP(Q,-A4/A1,Y,X)),
   IF FTEST(REDUCE(EQ,Y,X)) THEN RETURN(Q) ELSE RETURN(FALSE))$

ODE1(EQ,Y,X):=BLOCK([DE,F,G,Q],
   IF DERIVDEGREE(DE: EXPAND(LHS(EQ)-RHS(EQ)),Y,X) # 1
     THEN RETURN(FAILURE(MES1,EQ)),
   IF LINEAR2(DE,'DIFF(Y,X)) = FALSE THEN RETURN(FAILURE(MES2,EQ)),
   DE: SOLVE1(DE,'DIFF(Y,X)),
   IF FTEST(SOLVELNR(DE,Y,X)) THEN RETURN(Q),
   IF FTEST(INTFACTOR(G,F,Y,X)) THEN RETURN(EXACT(Q*G,Q*F,Y,X)),
                                 /* LINEAR2 binds F and G */
   IF FTEST(SOLVEHOM(DE,Y,X)) THEN RETURN(Q),
   IF FTEST(SOLVEBERNOULLI(DE,Y,X)) THEN RETURN(Q),
   IF FTEST(GENHOM(DE,Y,X)) THEN RETURN(Q) ELSE RETURN(FALSE))$

PR2(F):=FREEOF(Y,'DIFF(Y,X),'DIFF(Y,X,2),F)$

FTEST(CALL):=IS(NOT((Q: CALL)=FALSE))$

SOLVE1(EQ,Y):=
   BLOCK([DISPFLAG,EQ1],DISPFLAG:FALSE,EQ1:SOLVE(EQ,Y),FIRST(EV(EQ1)))$
```

```
SOLVEBERNOULLI(EQ,Y,X):=BLOCK([A1,A2,N],
   A1: COEFF(EQ: EXPAND(RHS(EQ)),Y,1),
   N: HIPOW(RATSIMP(EQ-A1*Y),Y),
   A2: COEFF(EQ,Y,N),
   IF NOT(NUMBERP(N)) OR N=0 OR NOT(EQ = A1*Y + A2*Y^N) THEN RETURN(FALSE),
   A1: INTEGRATE(A1,X),
   METHOD: 'BERNOULLI, ODEINDEX: N,
   RETURN(Y = %E^A1 * ((1-N)*INTEGRATE(A2*%E^((N-1)*A1),X) + 'C) ^ (1/(1-N))))$

/*  Generalized homogeneous equation:  y' = y/x * H(yx^n)
        Reference:  Moses' thesis.  */

GENHOM(EQ,Y,X):=BLOCK([G,U,N,A1,A2,A3],
   G: RHS(EQ)*X/Y,
   N: RATSIMP(X*DIFF(G,X)/(Y*DIFF(G,Y))),
   IF NOT(FREEOF(X,Y,N)) THEN RETURN(FALSE),
   A1: RATSIMP(SUBST(U/X^N,Y,G)),
   A2: INTEGRATE(1/(U*(N+A1)),U),
   A3: RATSIMP(SUBST(Y*X^N,U,A2)),
   METHOD: 'GENHOM, ODEINDEX: N,
   RETURN(X = 'C*%E^A3))$

/*  Chain of solution methods for second order linear homogeneous equations  */

HOM2(A1,A2,A3,Y,X):=
   IF FTEST(CC2(A2/A1,A3/A1,Y,X)) THEN Q ELSE
   IF FTEST(EXACT2(A1,A2,A3,Y,X)) THEN Q ELSE
   IF FTEST(XCC2(A1,A2,A3,Y,X)) THEN Q ELSE FALSE$

/*  B&D1P, pp. 106-112  */

CC2(F,G,Y,X):=BLOCK([A,SIGN,RADPRODEXPAND,ALPHA],
   IF NOT(FREEOF(X,Y,F) AND FREEOF(X,Y,G)) THEN RETURN(FALSE),
   METHOD: 'CONSTCOEFF, RADPRODEXPAND: FALSE,
   SIGN: ASKSIGN(A: F^2-4*G),
   IF SIGN = ZERO THEN RETURN(Y = %E^(-F*X/2) * ('K1 + 'K2*X)),
   IF SIGN = POS THEN
     RETURN(Y = 'K1*%E^((-F+SQRT(A))*X/2) + 'K2*%E^((-F-SQRT(A))*X/2)),
   A: -A, ALPHA: X*SQRT(A)/2,
   IF EXPONENTIALIZE = FALSE THEN
     RETURN(Y = %E^(-F*X/2) * ('K1*SIN(ALPHA) + 'K2*COS(ALPHA))),
   RETURN(Y = %E^(-F*X/2) * ('K1*EXP(%I*ALPHA) + 'K2*EXP(-%I*ALPHA))))$

/*  B&D1P, pp. 98-99, problem 17  */
```

```
SOLVE2(EQ,Y):=BLOCK([DISPFLAG,EQ1],
   DISPFLAG:FALSE,EQ1:SOLVE(EQ,Y),
   IF NOT(LENGTH(EQ1)=1) THEN RETURN(FAILURE(MES4,EV(EQ1))),
   FIRST(EV(EQ1)))$

MATCHDECLARE([F,G],FREEOF(X))$
DEFMATCH(LINEAR2,F*X+G,X)$

/*  B&D1P, pp. 13-14  */

SOLVELNR(EQ,Y,X):=BLOCK([F,G,W],
   IF LINEAR2(RHS(EQ),Y) = FALSE THEN RETURN(FALSE),
   W: %E^(INTEGRATE(F,X)),
   METHOD: 'LINEAR,
   RETURN(Y=W*(INTEGRATE(G/W,X)+'C)))$

/*  B&D1P, pp. 34-41  */

INTFACTOR(M,N,Y,X):=BLOCK([B1,B2,DMDX,DMDY,DNDX,DNDY,DD],
   DMDX: RATSIMP(DIFF(M,X)),  DMDY: RATSIMP(DIFF(M,Y)),
   DNDX: RATSIMP(DIFF(N,X)),  DNDY: RATSIMP(DIFF(N,Y)),
   IF (DD: DMDY-DNDX) = 0 THEN RETURN(1),
   IF DMDX-DNDY=0 AND DMDY+DNDX=0 THEN RETURN(1/(M^2 + N^2)),
   IF FREEOF(Y, (B1: RATSIMP(DD/N))) THEN RETURN(%E^(INTEGRATE(B1,X))),
   IF FREEOF(X, (B2: RATSIMP(DD/M)))
     THEN RETURN(%E^(INTEGRATE(-B2,Y))) ELSE RETURN(FALSE))$

EXACT(M,N,Y,X):=BLOCK([A,B],
   INTFACTOR: SUBST(YOLD,YNEW,Q),
   A: INTEGRATE(RATSIMP(M),X),
   B: RATSIMP(A + INTEGRATE(RATSIMP(N-DIFF(A,Y)),Y)),
   METHOD: 'EXACT,
   RETURN(B='C))$

/*  B&D1P, pp. 43-44  */

SOLVEHOM(EQ,Y,X):=BLOCK([QQ,A1,A2,A3],
   A1: RATSIMP(SUBST(X*QQ,Y,RHS(EQ))),
   IF NOT(FREEOF(X,A1)) THEN RETURN(FALSE),
   A2: INTEGRATE(1/(A1-QQ),QQ),
   A3: SUBST(Y/X,QQ,A2),
   METHOD: 'HOMOGENEOUS,
   RETURN(RATSIMP('C*X = %E^A3)))$

/*  B&D1P, p. 21, problem 15  */
```

```
EXACT2(A1,A2,A3,Y,X):=BLOCK([B1],
    IF DIFF(A1,X,2) - DIFF(A2,X) + A3 = 0
     THEN B1: %E^(-INTEGRATE((A2 - DIFF(A1,X))/A1, X))
     ELSE RETURN(FALSE),
    METHOD: 'EXACT,
    RETURN(Y = 'K1*B1*INTEGRATE(1/(A1*B1),X) + 'K2*B1))$

/*  B&D1P, pp. 113-114, problem 16  */

XCC2(A1,A2,A3,Y,X):=BLOCK([D,B1],
    IF A3=0 THEN RETURN(FALSE),
    D: RATSIMP((A1*DIFF(A3/A1,X) + 2*A2*A3/A1)/(2*(A3/A1)^(3/2))),
    IF FREEOF(X,Y,D) THEN B1: CC2(D,1,Y,Z) ELSE RETURN(FALSE),
    METHOD: 'XFORMTOCONSTCOEFF,
    RETURN(SUBST(INTEGRATE(SQRT(A3/A1),X),Z,B1)))$

/*  B&D1P, pp. 124-127  */

VARP(SOLN,G,Y,X):=BLOCK([Y1,Y2,Y3,Y4,WR],
    Y1: RATSIMP(SUBST(['K1=1,'K2=0],RHS(SOLN))),
    Y2: RATSIMP(SUBST(['K1=0,'K2=1],RHS(SOLN))),
    WR: Y1*DIFF(Y2,X) - Y2*DIFF(Y1,X),
    IF WR=0 THEN RETURN(FALSE),
    Y3: RATSIMP(Y1*G/WR),
    Y4: RATSIMP(Y2*G/WR),
    YP: RATSIMP(Y2*INTEGRATE(Y3,X) - Y1*INTEGRATE(Y4,X)),
    METHOD: 'VARIATIONOFPARAMETERS,
    RETURN(Y = RHS(SOLN) + YP))$

/*  Methods to reduce second-order equations free of x or y  */

REDUCE(EQ,Y,X):=BLOCK([B1,QQ],
    B1: SUBST(['DIFF(Y,X)=QQ, 'DIFF(Y,X,2)=QQ], EQ),
    IF FREEOF(Y,B1) THEN RETURN(NL1(EQ,Y,X)),
    IF FREEOF(X,B1) THEN RETURN(NL2(EQ,Y,X)) ELSE RETURN(FALSE))$

/*  B&D1P, p. 89, problem 1  */

NL1(EQ,Y,X):=BLOCK([DE,B,A1,A2,V],
    DE: SUBST(['DIFF(Y,X)=V, 'DIFF(Y,X,2)='DIFF(V,X)], EQ),
    IF (B: ODE1(DE,V,X)) = FALSE THEN RETURN(FALSE),
    A1: SUBST([V='DIFF(Y,X),'C='K1], B),
    A2: SOLVE2(A1,'DIFF(Y,X)),
    IF A2=FALSE THEN RETURN(FALSE),
    IF FTEST(ODE1(A2,Y,X))
```

APPENDIX II: Algorithm 407 DIFSUB

```
IMSL ROUTINE NAME   - DGEAR

PURPOSE             - DIFFERENTIAL EQUATION SOLVER - VARIABLE ORDER
                        ADAMS PREDICTOR CORRECTOR METHOD OR
                        GEARS METHOD

USAGE               - CALL DGEAR (N,FCN,FCNJ,X,H,Y,XEND,TOL,METH,
                        MITER,INDEX,IWK,WK,IER)

ARGUMENTS    N      - INPUT NUMBER OF FIRST-ORDER DIFFERENTIAL
                        EQUATIONS.
             FCN    - NAME OF SUBROUTINE FOR EVALUATING FUNCTIONS.
                        (INPUT)
                        THE SUBROUTINE ITSELF MUST ALSO BE PROVIDED
                          BY THE USER AND IT SHOULD BE OF THE
                          FOLLOWING FORM
                            SUBROUTINE FCN (N,X,Y,YPRIME)
                            REAL X,Y(N),YPRIME(N)
                                  .
                                  .
                                  .
                        FCN SHOULD EVALUATE YPRIME(1),...,YPRIME(N)
                          GIVEN N,X, AND Y(1),...,Y(N). YPRIME(I)
                          IS THE FIRST DERIVATIVE OF Y(I) WITH
                          RESPECT TO X.
                        FCN MUST APPEAR IN AN EXTERNAL STATEMENT IN
                          THE CALLING PROGRAM AND N,X,Y(1),...,Y(N)
                          MUST NOT BE ALTERED BY FCN.
             FCNJ   - NAME OF THE SUBROUTINE FOR COMPUTING THE
                        JACOBIAN MATRIX OF PARTIAL DERIVATIVES.
                        (INPUT)
                        THE SUBROUTINE ITSELF MUST ALSO BE PROVIDED
                          BY THE USER.
                        IF MITER=1 IT SHOULD BE OF THE FOLLOWING
                          FORM
                            SUBROUTINE FCNJ (N,X,Y,PD)
                            REAL X,Y(N),PD(N,N)
                                  .
                                  .
                        FCNJ MUST EVALUATE PD(I,J), THE PARTIAL
                          DERIVATIVE OF YPRIME(I) WITH RESPECT TO
                          Y(J), FOR I=1,N AND J=1,N.
                        IF MITER= -1 IT SHOULD BE OF THE FOLLOWING
                          FORM
                            SUBROUTINE FCNJ (N,X,Y,PD)
                            REAL X,Y(N),PD(1)
                                  .
                                  .
                        FCNJ MUST EVALUATE PD IN BAND STORAGE MODE.
                          THAT IS, PD(N*(J-I+NLC)+I) IS THE PARTIAL
                          DERIVATIVE OF YPRIME(I) WITH RESPECT TO
                          Y(J).  NLC IS THE NUMBER OF LOWER
                          CODIAGONALS FOR THE BAND MATRIX.
                        FCNJ MUST APPEAR IN AN EXTERNAL STATEMENT IN
                          THE CALLING PROGRAM AND N,X,Y(1),...,Y(N)
```

```
                MUST NOT BE ALTERED BY FCNJ.
              FCNJ IS USED ONLY IF MITER IS EQUAL TO
                1 OR -1. OTHERWISE A DUMMY ROUTINE CAN
                BE SUBSTITUTED. SEE REMARK 1.
X          - INDEPENDENT VARIABLE. (INPUT AND OUTPUT)
              ON INPUT, X SUPPLIES THE INITIAL VALUE
                AND IS USED ONLY ON THE FIRST CALL.
              ON OUTPUT, X IS REPLACED WITH THE CURRENT
                VALUE OF THE INDEPENDENT VARIABLE AT WHICH
                INTEGRATION HAS BEEN COMPLETED.
H          - INPUT/OUTPUT.
              ON INPUT, H CONTAINS THE NEXT STEP SIZE IN
                X. H IS USED ONLY ON THE FIRST CALL.
              ON OUTPUT, H CONTAINS THE STEP SIZE USED
                LAST, WHETHER SUCCESSFULLY OR NOT.
Y          - DEPENDENT VARIABLES, VECTOR OF LENGTH N.
              (INPUT AND OUTPUT)
              ON INPUT, Y(1),...,Y(N) SUPPLY INITIAL
                VALUES.
              ON OUTPUT, Y(1),...,Y(N) ARE REPLACED WITH
                A COMPUTED VALUE AT XEND.
XEND       - INPUT VALUE OF X AT WHICH SOLUTION IS DESIRED
              NEXT. INTEGRATION WILL NORMALLY GO
              BEYOND XEND AND THE ROUTINE WILL INTERPOLATE
              TO X = XEND.
            NOTE THAT (X-XEND)*H MUST BE LESS THAN
              ZERO (X AND H AS SPECIFIED ON INPUT).
TOL        - INPUT RELATIVE ERROR BOUND. TOL MUST BE
              GREATER THAN ZERO. TOL IS USED ONLY ON THE
              FIRST CALL UNLESS INDEX IS EQUAL TO -1.
              TOL SHOULD BE AT LEAST AN ORDER OF
              MAGNITUDE LARGER THAN THE UNIT ROUNDOFF
              BUT GENERALLY NOT LARGER THAN .001.
              SINGLE STEP ERROR ESTIMATES DIVIDED BY
              YMAX(I) WILL BE KEPT LESS THAN TOL IN
              ROOT-MEAN-SQUARE NORM (EUCLIDEAN NORM
              DIVIDED BY SQRT(N)). THE VECTOR YMAX OF
              WEIGHTS IS COMPUTED INTERNALLY AND STORED
              IN WORK VECTOR WK. INITIALLY YMAX(I) IS
              THE ABSOLUTE VALUE OF Y(I), WITH A DEFAULT
              VALUE OF ONE IF Y(I) IS EQUAL TO ZERO.
              THEREAFTER, YMAX(I) IS THE LARGEST VALUE
              OF THE ABSOLUTE VALUE OF Y(I) SEEN SO FAR,
              OR THE INITIAL VALUE OF YMAX(I) IF THAT IS
              LARGER.
METH       - INPUT BASIC METHOD INDICATOR.
              USED ONLY ON THE FIRST CALL UNLESS INDEX IS
              EQUAL TO -1.
            METH = 1, IMPLIES THAT THE ADAMS METHOD IS
              TO BE USED.
            METH = 2, IMPLIES THAT THE STIFF METHODS OF
              GEAR, OR THE BACKWARD DIFFERENTIATION
```

```
                FORMULAE ARE TO BE USED.
MITER   - INPUT ITERATION METHOD INDICATOR.
            MITER = 0, IMPLIES THAT FUNCTIONAL
              ITERATION IS USED. NO PARTIAL
              DERIVATIVES ARE NEEDED. A DUMMY FCNJ
              CAN BE USED.
            MITER = 1, IMPLIES THAT THE CHORD METHOD
              IS USED WITH AN ANALYTIC JACOBIAN. FOR
              THIS METHOD, THE USER SUPPLIES
              SUBROUTINE FCNJ.
            MITER = 2, IMPLIES THAT THE CHORD METHOD
              IS USED WITH THE JACOBIAN CALCULATED
              INTERNALLY BY FINITE DIFFERENCES.
              A DUMMY FCNJ CAN BE USED.
            MITER = 3, IMPLIES THAT THE CHORD METHOD
              IS USED WITH THE JACOBIAN REPLACED BY
              A DIAGONAL APPROXIMATION BASED ON A
              DIRECTIONAL DERIVATIVE.
              A DUMMY FCNJ CAN BE USED.
            MITER = -1 OR -2, IMPLIES USE THE SAME
              METHOD AS FOR MITER= 1 OR 2, RESPECTIVELY,
              BUT USING A BANDED JACOBIAN MATRIX.  IN
              THESE TWO CASES BANDWIDTH INFORMATION
              MUST BE PASSED TO DGEAR THROUGH THE
              COMMON BLOCK
                 COMMON /DBAND/ NLC,NUC
              WHERE NLC=NUMBER OF LOWER CODIAGONALS
                    NUC=NUMBER OF UPPER CODIAGONALS
INDEX   - INPUT AND OUTPUT PARAMETER USED TO INDICATE
            THE TYPE OF CALL TO THE SUBROUTINE.  ON
            OUTPUT INDEX IS RESET TO 0 IF INTEGRATION
            WAS SUCCESSFUL.  OTHERWISE, THE VALUE OF
            INDEX IS UNCHANGED.
          ON INPUT, INDEX = 1, IMPLIES THAT THIS IS THE
            FIRST CALL FOR THIS PROBLEM.
          ON INPUT, INDEX = 0, IMPLIES THAT THIS IS NOT
            THE FIRST CALL FOR THIS PROBLEM.
          ON INPUT, INDEX = -1, IMPLIES THAT THIS IS NOT
            THE FIRST CALL FOR THIS PROBLEM, AND THE
            USER HAS RESET TOL.
          ON INPUT, INDEX = 2, IMPLIES THAT THIS IS NOT
            THE FIRST CALL FOR THIS PROBLEM. INTEGRATION
            IS TO CONTINUE AND XEND IS TO BE HIT EXACTLY
            (NO INTERPOLATION IS DONE). THIS VALUE OF
            INDEX ASSUMES THAT XEND IS BEYOND THE
            CURRENT VALUE OF X.
          ON INPUT, INDEX = 3, IMPLIES THAT THIS IS NOT
            THE FIRST CALL FOR THIS PROBLEM. INTEGRATION
            IS TO CONTINUE AND CONTROL IS TO BE RETURNED
            TO THE CALLING PROGRAM AFTER ONE STEP. XEND
            IS IGNORED.
```

```
IWK      - INTEGER WORK VECTOR OF LENGTH N. USED ONLY IF
             MITER = 1,2,-1 OR -2.
WK       - REAL WORK VECTOR OF LENGTH 4*N+NMETH+NMITER.
             THE VALUE OF NMETH DEPENDS ON THE VALUE OF
               METH.
               IF METH IS EQUAL TO 1,
                 NMETH IS EQUAL TO N*13.
               IF METH IS EQUAL TO 2,
                 NMETH IS EQUAL TO N*6.
             THE VALUE OF NMITER DEPENDS ON THE VALUE OF
               MITER.
               IF MITER IS EQUAL TO 1 OR 2,
                 NMITER IS EQUAL TO N*(N+1)
               IF MITER IS EQUAL TO -1 OR -2,
                 NMITER IS EQUAL TO (2*NLC+NUC+3)*N
                  WHERE NLC=NUMBER OF LOWER CODIAGONALS
                        NUC=NUMBER OF UPPER CODIAGONALS
               IF MITER IS EQUAL TO 3,
                 NMITER IS EQUAL TO N.
               IF MITER IS EQUAL TO 0,
                 NMITER IS EQUAL TO 1.
             WK MUST REMAIN UNCHANGED BETWEEN SUCCESSIVE
             CALLS DURING INTEGRATION.
IER      - ERROR PARAMETER. (OUTPUT)
           WARNING ERROR
             IER = 33, IMPLIES THAT X+H WILL EQUAL X ON
               THE NEXT STEP. THIS CONDITION DOES NOT
               FORCE THE ROUTINE TO HALT. HOWEVER, IT
               DOES INDICATE ONE OF TWO CONDITIONS.
               THE USER MIGHT BE REQUIRING TOO MUCH
               ACCURACY VIA THE INPUT PARAMETER TOL.
               IN THIS CASE THE USER SHOULD CONSIDER
               INCREASING THE VALUE OF TOL. THE OTHER
               CONDITION WHICH MIGHT GIVE RISE TO THIS
               ERROR MESSAGE IS THAT THE SYSTEM OF
               DIFFERENTIAL EQUATIONS BEING SOLVED
               IS STIFF (EITHER IN GENERAL OR OVER
               THE SUBINTERVAL OF THE PROBLEM BEING
               SOLVED AT THE TIME OF THE ERROR). IN
               THIS CASE THE USER SHOULD CONSIDER
               USING A NONZERO VALUE FOR THE INPUT
               PARAMETER MITER.
           WARNING WITH FIX ERROR
             IER = 66, IMPLIES THAT THE ERROR TEST
               FAILED. H WAS REDUCED BY .1 ONE OR MORE
               TIMES AND THE STEP WAS TRIED AGAIN
               SUCCESSFULLY.
             IER = 67, IMPLIES THAT CORRECTOR
               CONVERGENCE COULD NOT BE ACHIEVED.
               H WAS REDUCED BY .1 ONE OR MORE TIMES AND
               THE STEP WAS TRIED AGAIN SUCCESSFULLY.
           TERMINAL ERROR
             IER = 132, IMPLIES THE INTEGRATION WAS
               HALTED AFTER FAILING TO PASS THE ERROR
```

```
                              TEST EVEN AFTER REDUCING H BY A FACTOR
                              OF 1.0E10 FROM ITS INITIAL VALUE.
                              SEE REMARKS.
                            IER = 133, IMPLIES THE INTEGRATION WAS
                              HALTED AFTER FAILING TO ACHIEVE
                              CORRECTOR CONVERGENCE EVEN AFTER
                              REDUCING H BY A FACTOR OF 1.0E10 FROM
                              ITS INITIAL VALUE. SEE REMARKS.
                            IER = 134, IMPLIES THAT AFTER SOME INITIAL
                              SUCCESS, THE INTEGRATION WAS HALTED EITHER
                              BY REPEATED ERROR TEST FAILURES OR BY
                              A TEST ON TOL. SEE REMARKS.
                            IER = 135, IMPLIES THAT ONE OF THE INPUT
                              PARAMETERS N,X,H,XEND,TOL,METH,MITER, OR
                              INDEX WAS SPECIFIED INCORRECTLY.
                            IER = 136, IMPLIES THAT INDEX HAD A VALUE
                              OF -1 ON INPUT, BUT THE DESIRED CHANGES
                              OF PARAMETERS WERE NOT IMPLEMENTED
                              BECAUSE XEND WAS NOT BEYOND X.
                              INTERPOLATION TO X = XEND WAS PERFORMED.
                              TO TRY AGAIN, SIMPLY CALL AGAIN WITH
                              INDEX EQUAL TO -1 AND A NEW VALUE FOR
                              XEND.

PRECISION/HARDWARE  - SINGLE AND DOUBLE/H32
                    - SINGLE/H36,H48,H60

REQD. IMSL ROUTINES - DGRCS,DGRIN,DGRPS,DGRST,LUDATF,LUELMF,LEQT1B,
                        UERTST,UGETIO

NOTATION            - INFORMATION ON SPECIAL NOTATION AND
                        CONVENTIONS IS AVAILABLE IN THE MANUAL
                        INTRODUCTION OR THROUGH IMSL ROUTINE UHELP

REMARKS  1.  THE EXTERNAL SUBROUTINE FCNJ IS USED ONLY WHEN
             INPUT PARAMETER MITER IS EQUAL TO 1 OR -1. OTHERWISE,
             A DUMMY FUNCTION CAN BE USED. THE DUMMY SUBROUTINE
             SHOULD BE OF THE FOLLOWING FORM
               SUBROUTINE FCNJ (N,X,Y,PD)
               INTEGER N
               REAL Y(N),PD(N,N),X
               RETURN
               END
         2.  AFTER THE INITIAL CALL, IF A NORMAL RETURN OCCURRED
             (IER=0) AND A NORMAL CONTINUATION IS DESIRED, SIMPLY
             RESET XEND AND CALL DGEAR AGAIN. ALL OTHER
             PARAMETERS WILL BE READY FOR THE NEXT CALL. A CHANGE
             OF PARAMETERS WITH INDEX EQUAL TO -1 CAN BE MADE
             AFTER EITHER A SUCCESSFUL OR AN UNSUCCESSFUL RETURN.
         3.  THE COMMON BLOCKS /DBAND/ AND /GEAR/ NEED TO BE
             PRESERVED BETWEEN CALLS TO DGEAR. IF IT IS NECESSARY
             FOR THE COMMON BLOCKS TO EXIST IN THE CALLING PROGRAM
```

THE FOLLOWING STATEMENTS SHOULD BE INCLUDED

```
  COMMON  /DBAND/ NLC,NUC
  COMMON  /GEAR/ DUMMY(48),SDUMMY(4),IDUMMY(38)
```

WHERE DUMMY, SDUMMY, AND IDUMMY ARE VARIABLE NAMES NOT USED ELSEWHERE IN THE CALLING PROGRAM. (FOR DOUBLE PRECISION DUMMY IS TYPE DOUBLE AND SDUMMY IS TYPE REAL)

4. THE CHOICE OF VALUES FOR METH AND MITER MAY REQUIRE SOME EXPERIMENTATION, AND ALSO SOME CONSIDERATION OF THE NATURE OF THE PROBLEM AND OF STORAGE REQUIREMENTS. THE PRIME CONSIDERATION IS STIFFNESS. IF THE PROBLEM IS NOT STIFF, THE BEST CHOICE IS PROBABLY METH = 1 WITH MITER = 0. IF THE PROBLEM IS STIFF TO A SIGNIFICANT DEGREE, THEN METH SHOULD BE 2 AND MITER SHOULD BE 1,2,-1,-2 OR 3. IF THE USER HAS NO KNOWLEDGE OF THE INHERENT TIME CONSTANTS OF THE PROBLEM, WITH WHICH TO PREDICT ITS STIFFNESS, ONE WAY TO DETERMINE THIS IS TO TRY METH = 1 AND MITER = 0 FIRST, AND LOOK AT THE BEHAVIOR OF THE SOLUTION COMPUTED AND THE STEP SIZES USED. IF THE TYPICAL VALUES OF H ARE MUCH SMALLER THAN THE SOLUTION BEHAVIOR WOULD SEEM TO REQUIRE (THAT IS, MORE THAN 100 STEPS ARE TAKEN OVER AN INTERVAL IN WHICH THE SOLUTIONS CHANGE BY LESS THAN ONE PERCENT), THEN THE PROBLEM IS PROBABLY STIFF AND THE DEGREE OF STIFFNESS CAN BE ESTIMATED FROM THE VALUES OF H USED AND THE SMOOTHNESS OF THE SOLUTION. IF THE DEGREE OF STIFFNESS IS ONLY SLIGHT, IT MAY BE THAT METH=1 IS MORE EFFICIENT THAN METH=2. EXPERIMENTATION WOULD BE REQUIRED TO DETERMINE THIS. REGARDLESS OF METH, THE LEAST EFFECTIVE VALUE OF MITER IS 0, AND THE MOST EFFECTIVE IS 1,-1,2,OR -2. MITER = 3 IS GENERALLY SOMEWHERE IN BETWEEN. SINCE THE STORAGE REQUIREMENTS GO UP IN THE SAME ORDER AS EFFECTIVENESS, TRADE-OFF CONSIDERATIONS ARE NECESSARY. FOR REASONS OF ACCURACY AND SPEED, THE CHOICE OF ABS(MITER)=1 IS GENERALLY PREFERRED TO ABS(MITER)=2, UNLESS THE SYSTEM IS FAIRLY COMPLICATED (AND FCNJ IS THUS NOT FEASIBLE TO CODE). THE ACCURACY OF THE FCNJ CALCULATION CAN BE CHECKED BY COMPARISON OF THE JACOBIAN WITH THAT GENERATED WITH ABS(MITER)=2. IF THE JACOBIAN MATRIX IS SIGNIFICANTLY DIAGONALLY DOMINANT, THEN THE OPTION MITER = 3 IS LIKELY TO BE NEARLY AS EFFECTIVE AS ABS(MITER)=1 OR 2, AND WILL SAVE CONSIDERABLE STORAGE AND RUN TIME. IT IS POSSIBLE, AND POTENTIALLY QUITE DESIRABLE, TO USE DIFFERENT VALUES OF METH AND MITER IN DIFFERENT SUBINTERVALS OF THE PROBLEM. FOR EXAMPLE, IF THE PROBLEM IS NON-STIFF INITIALLY AND STIFF LATER, METH = 1 AND MITER = 0 MIGHT BE SET INITIALLY, AND METH = 2 AND MITER = 1 LATER.

5. THE INITIAL VALUE OF THE STEP SIZE, H, SHOULD BE CHOSEN CONSIDERABLY SMALLER THAN THE AVERAGE VALUE EXPECTED FOR THE PROBLEM, AS THE FIRST-ORDER METHOD

WITH WHICH DGEAR BEGINS IS NOT GENERALLY THE MOST EFFICIENT ONE. HOWEVER, FOR THE FIRST STEP, AS FOR EVERY STEP, DGEAR TESTS FOR THE POSSIBILITY THAT THE STEP SIZE WAS TOO LARGE TO PASS THE ERROR TEST (BASED ON TOL), AND IF SO ADJUSTS THE STEP SIZE DOWN AUTOMATICALLY. THIS DOWNWARD ADJUSTMENT, IF ANY, IS NOTED BY IER HAVING THE VALUES 66 OR 67, AND SUBSEQUENT RUNS ON THE SAME OR SIMILAR PROBLEM SHOULD BE STARTED WITH AN APPROPRIATELY SMALLER VALUE OF H.

6. SOME OF THE VALUES OF INTEREST LOCATED IN THE COMMON BLOCK /GEAR/ ARE
   A. HUSED, THE STEP SIZE H LAST USED SUCCESSFULLY (DUMMY(8))
   B. NQUSED, THE ORDER LAST USED SUCCESSFULLY (IDUMMY(6))
   C. NSTEP, THE CUMULATIVE NUMBER OF STEPS TAKEN (IDUMMY(7))
   D. NFE, THE CUMULATIVE NUMBER OF FCN EVALUATIONS (IDUMMY(8))
   E. NJE, THE CUMULATIVE NUMBER OF JACOBIAN EVALUATIONS, AND HENCE ALSO OF MATRIX LU DECOMPOSITIONS (IDUMMY(9))
7. THE NORMAL USAGE OF DGEAR MAY BE SUMMARIZED AS FOLLOWS
   A. SET THE INITIAL VALUES IN Y.
   B. SET N, X, H, TOL, METH, AND MITER.
   C. SET XEND TO THE FIRST OUTPUT POINT, AND INDEX TO 1.
   D. CALL DGEAR
   E. EXIT IF IER IS GREATER THAN 128.
   F. OTHERWISE, DO DESIRED OUTPUT OF Y.
   G. EXIT IF THE PROBLEM IS FINISHED.
   H. OTHERWISE, RESET XEND TO THE NEXT OUTPUT POINT, AND RETURN TO STEP D.
8. THE ERROR WHICH IS CONTROLLED BY WAY OF THE PARAMETER TOL IS AN ESTIMATE OF THE LOCAL TRUNCATION ERROR, THAT IS, THE ERROR COMMITTED ON TAKING A SINGLE STEP WITH THE METHOD, STARTING WITH DATA REGARDED AS EXACT. THIS IS TO BE DISTINGUISHED FROM THE GLOBAL TRUNCATION ERROR, WHICH IS THE ERROR IN ANY GIVEN COMPUTED VALUE OF Y(X) AS A RESULT OF THE LOCAL TRUNCATION ERRORS FROM ALL STEPS TAKEN TO OBTAIN Y(X). THE LATTER ERROR ACCUMULATES IN A NON-TRIVIAL WAY FROM THE LOCAL ERRORS, AND IS NEITHER ESTIMATED NOR CONTROLLED BY THE ROUTINE. SINCE IT IS USUALLY THE GLOBAL ERROR THAT A USER WANTS TO HAVE UNDER CONTROL, SOME EXPERIMENTATION MAY BE NECESSARY TO GET THE RIGHT VALUE OF TOL TO ACHIEVE THE USERS NEEDS. IF THE PROBLEM IS MATHEMATICALLY STABLE, AND THE METHOD USED IS APPROPRIATELY STABLE, THEN THE GLOBAL ERROR AT A GIVEN X SHOULD VARY SMOOTHLY WITH TOL IN A MONOTONE INCREASING MANNER.

9. IF THE ROUTINE RETURNS WITH IER VALUES OF 132, 133, OR 134, THE USER SHOULD CHECK TO SEE IF TOO MUCH ACCURACY IS BEING REQUIRED. THE USER MAY WISH TO SET TOL TO A LARGER VALUE AND CONTINUE. ANOTHER POSSIBLE CAUSE OF THESE ERROR CONDITIONS IS AN ERROR IN THE CODING OF THE EXTERNAL FUNCTIONS FCN OR FCNJ. IF NO ERRORS ARE FOUND, IT MAY BE NECESSARY TO MONITOR INTERMEDIATE QUANTITIES GENERATED BY THE ROUTINE. THESE QUANTITIES ARE STORED IN THE WORK VECTOR WK AND INDEXED BY SPECIFIC ELEMENTS IN THE COMMON BLOCK /GEAR/. IF IER IS 132 OR 134, THE COMPONENTS CAUSING THE ERROR TEST FAILURE CAN BE IDENTIFIED FROM LARGE VALUES OF THE QUANTITY
   WK(IDUMMY(11)+I)/WK(I), FOR I=1,...,N.
   ONE CAUSE OF THIS MAY BE A VERY SMALL BUT NONZERO INITIAL VALUE OF ABS(Y(I)).
   IF IER IS 133, SEVERAL POSSIBILITIES EXIST.
   IT MAY BE INSTRUCTIVE TO TRY DIFFERENT VALUES OF MITER. ALTERNATIVELY, THE USER MIGHT MONITOR SUCCESSIVE CORRECTOR ITERATES CONTAINED IN WK(IDUMMY(12)+I), FOR I=1,...,N. ANOTHER POSSIBILITY MIGHT BE TO MONITOR THE JACOBIAN MATRIX, IF ONE IS USED, STORED, BY COLUMN, IN WK(IDUMMY(10)+I), FOR I=1,...,N*N IF ABS(MITER) IS EQUAL TO 1 OR 2, OR FOR I=1,...,N IF MITER IS EQUAL TO 3.

Algorithm

DGEAR finds approximations to the solution of a system of first order ordinary differential equations of the form y'=f(x,y) with initial conditions. The basic methods used for the solution are of implicit linear multistep type. There are two classes of such methods available to the user. The first is the implicit Adams methods (up to order twelve), and the second is the backward differentiation formula (BDF) methods (up to order five), also called Gear's stiff methods. In either case the implicitness of the basic formula requires that an algebraic system of equations be solved at each step. A variety of corrector iteration methods is available for this.

DGEAR and the associated nuclei are adaptations of a package designed by A. C. Hindmarsh based on C. W. Gear's subroutine DIFSUB.

See references:

1. Hindmarsh, A.C., "GEAR: Ordinary Differential Equation System Solver", Lawrence Livermore Laboratory, Report UCID-30001, Revision 3, December, 1974.

2. Gear, C.W., Numerical Initial Value Problems in Ordinary Differential Equations, Prentice-Hall, Englewood Cliffs, New Jersey, 1971.

Example

This example illustrates the basic usage of DGEAR. A table of solution values for x = 1.0,2.0,...,10.0 is obtained for the predator-prey problem:

$$y_1' = 2y_1(1-y_2) \qquad y_1 = 1$$
$$\qquad\qquad\qquad\qquad\qquad\qquad \text{at } x = 0$$
$$y_2' = y_2(y_1-1) \qquad y_2 = 3$$

Input:

```
      INTEGER   N,METH,MITER,INDEX,IWK(2),IER,K
      REAL      Y(2),WK(35),X,TOL,XEND,H
      EXTERNAL  FCN,FCNJ
      N    = 2
      X    = 0.0
      Y(1) = 1.0
      Y(2) = 3.0
      TOL  = .00001
      H    = .00001
      METH =  1
      MITER=  0
      INDEX=  1
      DO 10 K=1,10
        XEND=FLOAT(K)
        CALL DGEAR (N,FCN,FCNJ,X,H,Y,XEND,TOL,METH,MITER,INDEX,IWK,WK,IER)
        IF(IER.GT.128) GO TO 20
C                     Y(1) and Y(2) are current solution values at X.
C                     Insert write statement here.
   10 CONTINUE
      STOP
   20 CONTINUE
C                     Handle IER.GT.128
C                     Items that may help diagnose the problem should be
C                     output here.
C                     TOL,N,Y(1),...,Y(N),XEND,H,X,METH,MITER, AND INDEX.
      STOP
      END
      SUBROUTINE     FCN(N,X,Y,YPRIME)
      INTEGER        N
      REAL           Y(N),YPRIME(N),X
      YPRIME(1) = 2.0*Y(1)*(1.0-Y(2))
      YPRIME(2) = Y(2)*(Y(1)-1.0)
      RETURN
      END

      SUBROUTINE FCNJ(N,X,Y,PD)
      INTEGER    N
      REAL       Y(N),PD(N,N),X
      RETURN
      END
```

Output:

IER = 0

X	Y(1)	Y(2)
1.	0.08	1.46
2.	0.085	0.58
3.	0.29	.25
4.	1.45	0.19

5.	4.05	1.44
6.	0.18	2.26
7.	0.07	0.91
8.	0.15	0.37
9.	0.65	0.19
10.	3.15	0.35

BIBLIOGRAPHY

Acton, J.P. and B.M. MItchell (1979). Evaluating time-of-day electricity rates for residential customers. *Rand report* R-2509-EWP.

Afriat, S.N. (1957). Orthogonal and oblique projectors and the characterestics of pairs of vector spaces. *Proc of the Cambridge Phil Soc* 53(4), 800-817.

__________ (1972). The theory of international comparisons of real incomes and prices. In *International Comparisons of Prices and Output,* edited by D.J.Daly (Studies in Income and Wealth no. 37; proceedings of the conference at York University, Toronto, 1970). New York: National Bureau of Economic Research, 13-84.

__________ (1977). *The Price Index.* Cambridge University Press.

__________ (1978). On Wald's 'New Formula' for the cost of living. In *Theory and Applications Of Economic Indices*, edited by W.Eichorn, R.Henn, O.Opitz, R.W.Shephard (Proceedings of an International Seminar held at the University of Karlsruhe, 23 June-1 July, 1976). Würsberg: Physica-Verlag.

__________ (1980). *Demand Functions and the Slutsky Matrix.* Princeton Studies in Mathematical Economics no. 7. Princeton University Press.

__________ (1987) *Logic of Choice and Economic Theory.* Clarendon Press, Oxford.

Aigner, D.J. (1984). The welfare econometrics of peak load pricing for electricity. *Journal of Econometrics* (26) 1-15.

Antonelli, G.B. (1886). *Sulla teoriamatematica della economia politica,* Pisa: Nella Tipografia del Fochetto, (trans. chapter 16 of J.S.Chipman, et.al., 1971).

Apostol, T.M. (1967). Calculus. Waltham, Mass: Xerox

Atkinson, S.E. (1983). The implications of homothetic separability for share equation price elasticities. *Journal of Business and Economic Statistics* (1) 211-14.

Balasko, Y. (1988). *Foundations of the Theory of General Equilibrium.* Academic Press, London.

Barnett,W.A. (1983). The recent appearance of the homotheticity restriction on preferences. *Journal of Business and Economic Statistics* (1) 215-218.

Barten, A.P. (1968). Estimating Demand Equations. *Econometrica,* (36) 213-22.

Bentley, D. and D. Cooke, (1973). *Linear Algebra and Differential Equations*. Holt, Reinhart and Winston.

Berge, C. (1963). *Topological Spaces* Oliver and Boyd, Edinburgh.

Bergson, A. (1975). A note on consumer's surplus. *Journal of Economic Literature* (13) 42-51.

Bishop, R.L. and S.L. Goldberg, (1968). *Tensor Analysis on Manifolds*. Dover Publications, New York.

Blackorby, C., D.Primont, and R.R.Russel. (1978). *Duality, Separability, and Functional Structure: Theory and Economic Applications*. Elsevier/North Holland, New York.

Blackorby, C. and W.E. Diewert. (1979). Expenditure functions, local duality, and second order approximations. *Econometrica* (47) 579-601.

Blackorby, C. D.Nisson, and R.R.Russell. (1970). Homothetic separability and consumer budgeting, *Econometrica* (38) 468-472.

Boadway, R. and N. Bruce (1984). *Welfare Economics*. Basil Blackwell, New York.

Boyce W.E. and DiPrima (1977). *Elementary Differential Equations and Boundary Value Problems,* (3rd Edition). John Wiley and Sons, New York.

Burns, M.E. (1973). A note on the concept and measure of consumer surplus. *Americal Economic Review* (53), 335-44.

———— (1977). On the uniqueness of consumer surplus and the invariance of index numbers. *The Manchester School of Economics and Social Studies* (45) 41-61.

Cartan, H. (1970). *Differential Forms*. Hermann, Paris.

Caves, D.W. and L.R. Christensen, (1980). Econometric analysis of residential time-of-use electricity pricing experiments. *Journal of Econometrics,* (14), 422-32.

Caves, D.W., L.R. Christensen, and J.A. Herriges (1984a). Consistency of residential response in time-of-use electricity pricing experiments. *Journal of Econometrics* (26), 179-203.

Caves, D.W., L.R. Christensen, P.E. Schoech, and W. Hendricks. (1984). A comparison of different methodologies in a case study of residential time-of-use electricity pricing. *Journal of Econometrics* (26), 17-33.

Chevally, C. (1946). *Theory of Lie Groups*. Princeton Mathematical Series, no. 22. Princeton University Press.

Chipman, J.S. and J. Moore (1976). The scope of consumer surplus arguments. In *Evolution, Welfare and Time in Economics: Essays in Honor of Nicholas Georgescu-Roegen,* edited by A.M.Tang; Lexington Mass.: Heath-Lexington Books 69-123.

——————— (1980). Compensating variation, consumer surplus and welfare. *American Economic Review* (70), 933-49.

Choquet-Bruhat, Y., C. DeWitt-Morette, and M. Dillard-Bleick (1982) *Analysis, Manifolds and Physics*. (Revised Edition) North-Holland, New York.

Collatz, L. (1960). *The Numerical Treatment of Differential Equations* (3rd edition). Springer Verlag, Berlin and New York.

Crew, M. (Ed.) (1980). *Issues in Public Utility Pricing of Electricity*. Lexington Books.

Deaton, A.S. (1978). Specification and testing in applied demand analysis. *Economic Journal* (88), 524-36.

——————— (1981) (ed). *Essays in Theory and Measurement of Demand: Essays in Honor of Sir Richard Stone*. Cambridge University Press, Cambridge and New York.

Diewert, W.E. (1974). Applications of duality theory. In *Frontiers of Econometrics,* Vol. 2. edited by Intrilligator, and D.Kendrick. North-Holland, Amsterdam.

Debreu, G. (1959). *Theory of Value*. Yale University Press. New Haven.

Debreu, G. (1972). Smooth Preferences. *Econometrica* (40), 603-15.

Debreu, G. (1976). Regular differentiable economies. *American Economic Review* (66) 2280-287.

Debreu, G. (1983). *Mathematical Economics*. (Translated by S.Devarajan) Cambridge University Press, Cambridge.

Dixit, A.K. and P.A.Weller (1979). The three consumer's surplus. *Economica* (46), 125-35.

Dupuit, J. (1844). De la mesure de l'utilité des travaux publics, Trans. in K.J.Arrow and T.Scitovsky (eds), *Readings in Welfare Economics*. Irwin, Homewood, Illinois.

Eggleston, H.G. (1958). *Convexity* (2nd edition). Cambridge University Press.

Fenchel, W. (1949). On Conjugate Convex Functions. *Canadian Journal of Mathematics* (1), 73-7.

Foster, C.D. and H.L.I. Neuberger, (1974). The ambiguity of the consumer's surplus measure of welfare change. *Oxford Economic Papers* (26), 66-77.

Gear, C.W. (1971a). *Numerical Initial Value Problems in Differential Equations*. Prentice Hall, Englewood Cliffs New Jersey.

———— (1971b). Algorithm 407 DIFSUB for solution of ordinary differential equations. *Automation and Computing Machinery* (14), 185-90.

Goldberger, A.S. (1987). *Functional Form and Utility: A Review of Consumer Demand Theory*. Westview Press, London.

Golden, J.P. (1977). MACSYMA's symbolic ODE solver. *Proceedings of the 1977 MACSYMA Users Conference,* U.S. Dept of Commerce.

Goldman, S.M. and H. Uzawa, (1964). A note on separability in demand analysis. *Econometrica* (32), 387-98.

Gorman, W.M. (1961). On a class of preference fields. *Metroeconomica* (13), 53-6.

———— (1968). The structure of utility functions. *Review of Economic Studies* 35(4), 367-90.

———— (1976). Tricks with utility functions. In *Essays in Economic Analysis*, edited by M.J.Artis and A.R.Nobay. Cambridge University Press.

Harberger, A.C. (1971). Three basic postulates for applied welfare economics: An interpretative essay. *Journal Of Economic Literature* (9), 785-97.

Hartman, P. (1964). *Ordinary Differential Equations*. Wiley, New York.

Hause, J.C. (1975). The theory of welfare measurement. *Journal of* Political Economy (83), 1154-78.

Hausman, J.A., M. Kinnucan, and D. McFadden, (1979). A two level electricity demand model. *Journal of Econometrics* (9), 263-89.

Hausman, J.A. (1981). Exact consumer surplus and deadweight loss. *American Economic Review* (71), 662-676.

Henrici, P. (1967). *Elements Of Numerical Analysis*. John Wiley and Sons, New York.

Hicks, J.R. (1943). The four consumer's surpluses. *Review of Economic Studies,* (49), 696-712.

———— (1956). *A Revision of Demand Theory*. Clarendon Press, Oxford.

Hicks, N.J. (1964). *Notes on Differential Geometry*. Van Nostrand, London.

Hirsch, M. (1976). *Differentiable Topology*. Springer-Verlag, New York.

Hotelling, H. (1938). The general welfare in relation to problems of taxation and of railway and utility rates. *Econometrica,* (6), 242-69.

Howrey, E.P. and H.R. Varian (1984). Estimating the distributional impact of time-of-day pricing of electricity. *Journal of Econometrics,* (26), 65-82.

Hurwicz, L. (1971). On the problem of integrability of demand functions, in *Preferences, Utility and Demand,* edited by Chipman,

J.S., L.Hurwicz, M.K.Ritcher, and H.F.Sonnenschein. Harcourt Brace Jovanovich, New York.

Hurwicz, L. and H. Uzawa (1971). On the integrability of demand functions, in Chipman et al (ibid).

Israel, A.B. (1974). *Generalized Inverses*. John Wiley and Sons.

Just, R.E., D.L.Heuth, and A. Schmitz, (1982). *Applied Welfare Economics and Welfare Policy*. Prentice-Hall, Englewood Cliffs, New Jersey.

Katzner, D.W. (1970). *Static Demand Theory*. Macmillan Company, New York.

King, M.A. (1983). Welfare analysis of tax reforms using household data. *Journal of Public Economics* (21), 183-214.

Klein, L. and H. Rubin, (1947). A constant utility index of the cost of living. *Review of Economic Studies* (15), 84-7.

Kohler, D.F. (1983). The bias in price elasticity estimates under homothetic separability: implications for analysis of peak load electricity pricing. *Journal of Business and Economic Statistics* (1), 202-10.

Kohler, D.F. and B.M. Mitchell (1984). Response to residential Time-of-Use Electricity Rates. How transferable are the findings? *Journal of Econometrics*. (26) 141-77.

Lawrence, A. and S. Braithwaite, (1979). The residential demand for electricity with time-of-day pricing. *Journal Of Econometrics* (9), 59-77.

Malko, J. and A. Faruqi. (1979). Implementing Time-of-Day pricing of electricity: some current challeges and activities. In M. Crew (1980).

Mas-Colell, A. (1985). *The Theory of General Equilibrium. A Differential Approach*. Cambridge University Press, Cambridge.

Marshall, A. (1890). *Principles of Economics*. Macmillan, London.

Matsushima, Y. (1972). *Differentiable Manifolds*. Marcel Decker, New York.

McKenzie, G. and I. Pearce, (1976). Exact measures of welfare and the cost of living. *Review of Economic Studies* (135), 465-8.

McKenzie, G. (1983). *Measuring Economic Welfare: New Methods*. Cambridge University Press.

McKenzie, L.W. (1957). Demand theory without a utility index. *Review of Economic Studies,* (24), 185-9.

Mishan, E.J. (1976). *Cost Benefit Analysis* (2nd ed). Praeger Publishers, New York.

Mitchell, B., W.G. Manning, and J. Acton,(1978). *Peak Load Pricing*. Rand Corporation Research Study. Ballinger Publishing Company.

Mitchell, B.M. and J.P.Acton. (1983). Electricity consumption by time-of-Usein a hybrid demand system, in J. Finsinger (ed.) *Public Sector Economics*. Macmillan, London.

Mohring, H. (1971). Alternative welfare gain and loss measures. *Western Economic Journal,* (9), 349-68.

Moses, J. (1971). Symbolic Integration: The stormy decade. *Automation and Computer Machines* (14), 548-60.

Murty, K.N. (1982). Theoretical restrictions on the parameters of indirect addilog demand equations - a comment. *Econometrica* (50), 225-6.

Newman, P. (1969). Some properties of concave functions. *Journal Of Economic Theory,* (1), 291-34.

Parks, R.E. (1983). Comment. *Journal of Business and Economic Statistics*. (1) 221-226.

Parks, R.W. (1969). Systems of demand equations: An empirical comparison of alternative functional forms, *Econometrica* (37) 629-650.

Parks, R.W. and D. Weitzel, (1984). Measuring the consumer welfare effects of time differentiated electricity prices. *Journal of Econometrics* (26), 35-64.

Pollak, R.A. (1971). Additive utility functions and linear Engel curves. *Review Of Economic Studies* (38), 401-14.

Phlips, L. (1974). *Applied Consumption Analysis*. North-Holland, Amsterdam.

Rader, T. (1976). Equivalence of consumer surplus, the divisia index of output, and Eisenberg's addilog social utility. *Journal Of Theory,* (20), 271-75.

Rockafeller, R.T. (1970). *Convex Analysis,* Princeton University Press.

Samuelson, P.A. (1942). Constancy of the marginal utility of income, in *Studies in Mathematical Economics and Econometrics: in Memory of Henry Schultz,* edited by Lange, O. University of Chicago Press, Chicago.

———— (1947). *Foundations Of Economic Analysis,* Harvard University Press, Cambridge, Mass.

——— (1950). The problem of integrability in utility theory *Economica,* (17), 355-85.

Seade, J. (1978). Consumer surplus and the linearity of Engel curves. *Economic Journal* (88), 411-523.

Shephard, R. (1953). *Cost and Production Functions.* Princeton University Press.

Silberberg, E. (1972). Duality and the many consumer surplus. *American Economic Review* (62), 942-52.

Spivak, M. (1979). *A comprehensive Introduction to Differential Geometry,* 5 volumes (second edition) Publish or Perish, Boston.

Sternberg, S. (1964). *Lectures on Differential Geometry.* Prentice Hall, Englewood Cliffs, New Jersey.

Stone, R.N. (1954). Linear Expenditure Systems and demand analysis: an application to the pattern of British demand. *Economic Journal* (64), 511-27.

Taylor, L. (1975). The demand for electricity: A survey. *Bell Journal of Economics,* (6), 110-32.

Turvey, R. (1974). How to judge when price changes will improve resource allocation. *Economic Journal* 825-32.

Vartia, Y.O. (1983). Efficient methods of measuring welfare change and compensated income in terms of ordinary demand functions. *Econometrica* (51), 79-97.

Wenders, J.T. and L.D. Taylor, (1976). Experiments in seasonal time-of-day pricing of electricity for residential users. *Bell Journal of Economics,* (7).

Willig, R.D. (1976). Consumer's surplus without apology. *American Economic Review* (66), 589-97.

——— (1973). *Consumer's Surplus: A Rigorous Cookbook,* Technical Report No. 98, Stanford University Press.

——— (1979). Consumer's surplus without apology: a reply. *American Economic Review,* (69), 469-74.

Yoshihara, K. (1969). Demand functions: an application to the Japanese expenditure pattern. *Econometrica* (37), 257-74.

Zajac, E.E. (1979). Dupuit - Marshall consumer's surplus, utility and revealed preference. *Journal Of Economic Theory* (20), 260-70.